Am J Cancer Res 2020;10(12):4513-4526
www.ajcr.us /ISSN:2156-6976/ajcr0121568

Original Article

Prediction of KRAS, NRAS and BRAF status in colorectal cancer patients with liver metastasis using a deep artificial neural network based on radiomics and semantic features

Ruichuan Shi[1,2,3,4*], Weixing Chen[5*], Bowen Yang[1,2,3,4], Jinglei Qu[1,2,3,4], Yu Cheng[1,2,3,4], Zhitu Zhu[6], Yu Gao[6], Qian Wang[7], Yunpeng Liu[1,2,3,4], Zhi Li[1,2,3,4], Xiujuan Qu[1,2,3,4]

[1]Department of Medical Oncology, The First Hospital of China Medical University, 110001, Liaoning, China; [2]Key Laboratory of Anticancer Drugs and Biotherapy of Liaoning Province, The First Hospital of China Medical University, 110001, Liaoning, China; [3]Liaoning Province Clinical Research Center for Cancer, 110001, Liaoning, China; [4]Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors, Ministry of Education, 110001, Liaoning, China; [5]Paul C. Lauterbur Research Center for Biomedical Imaging, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 518005, Guangdong, China; [6]Cancer Center, The First Affiliated Hospital of Jinzhou Medical University, 121001, Liaoning, China; [7]Department of Medical Oncology, Liaoning Cancer Hospital and Institute, Cancer Hospital of China Medical University, 110042, Liaoning, China. [*]Equal contributors.

Received September 1, 2020; Accepted November 17, 2020; Epub December 1, 2020; Published December 15, 2020

Abstract: There is a critical need for development of improved methods capable of accurately predicting the RAS (KRAS and NRAS) and BRAF gene mutation status in patients with advanced colorectal cancer (CRC). The purpose of this study was to investigate whether radiomics and/or semantic features could improve the detection accuracy of RAS/BRAF gene mutation status in patients with colorectal liver metastasis (CRLM). In this retrospective study, 159 patients who had been diagnosed with CRLM in two hospitals were enrolled. All patients received lung and abdominal contrast-enhanced CT (CECT) scans prior to radiation therapy and chemotherapy. Semantic features were independently assessed by two radiologists. Radiomics features were extracted from the portal venous phase (PVP) of the CT scan for each patient. Seven machine learning algorithms were used to establish three scores based on the semantic, radiomics and the combination of both features. Two semantic and 851 radiomics features were used to predict the mutation status of RAS and BRAF using an artificial neural network method (ANN). This approach performed best out of the seven tested algorithms. We constructed three scores which were based on radiomics, semantic features and the combined scores. The combined score could distinguish between wild-type and mutant patients with an AUC of 0.95 in the primary cohort and 0.79 in the validation cohort. This study proved that the application of radiomics together with semantic features can improve non-invasive assessment of the gene mutation status of RAS (KRAS and NRAS) and BRAF in CRLM.

Keywords: RAS, BRAF, colorectal cancer, radiomics, artificial neural network

Introduction

Colorectal cancer (CRC) is the third most common and deadly tumor in males and females worldwide [1]. According to the World Health Organization (WHO), 1.8 million new cases of CRC were diagnosed worldwide in 2018, with almost 861,000 deaths. In recent years, death rates from CRC have progressively declined in many countries due to improved early detection and more effective primary and adjuvant therapies [2, 3]. However, the overall survival (OS) in patients with advanced CRC remains poor. Liver metastasis is the main cause of death in patients with CRC [4]. 15%-25% of patients are commonly diagnosed with synchronous liver metastases at the time of primary diagnosis [5].

Several studies have reported that patient outcomes can be improved through precise identification and stratification of patients harboring

specific mutant genes [6-8]. Therefore, genomic characterization to optimize patient selection is of great research and clinical interest. With improvements in precision medicine, RAS and BRAF gene mutation status should be used to inform decision making in the clinic [9-11]. Gain-of-function mutations of RAS and BRAF genes resulting in continuous activation of the RAS-mitogen-activated protein kinase (MAPK) pathway have been characterized in most cases of CRC. KRAS mutations represent an early event in CRC tumorigenesis and occur in 35-45% of CRC cases. BRAF mutations, prevalently c.1799T > A (V600E) mutation have been found in 4.7% to 8.7% of CRC [12]. For surgical treatment of these patients, it has been suggested that a resection margin of at least 10 mm should be obtained [13, 14]. However, there is no clear recommendation for the most optimal resection margin for CRLM patients harboring RAS and/or BRAF mutation as patients with both mutations have a higher risk of positive surgical margins [15]. Therefore, the selection of resection margin should be different for CRLM patients harboring RAS and/or BRAF mutation [16, 17]. In addition, RAS and BRAF mutations are associated with shorter disease-free survival (DFS) and OS [18-20]. These data suggest that pre-treatment CT scans could potentially be used to predict the RAS and BRAF mutation status in patients with CRLM and to further guide treatments with surgery and/or chemotherapy.

The detection of driver gene mutation status requires the analysis of specimens obtained from surgery or biopsy, which are invasive and expensive procedures. Since Lambin and his colleagues, first proposed the concept of Radiomics in 2012, this high-throughput, non-invasive strategy has been shown to provide addition information that can benefit clinical decision making in multi-centric settings and many different cancers [21-32]. Radiomics holds great promises in representing the underlying differences in the molecular phenotype of tumors [33]. Most previous radiomics studies have focused on KRAS mutation status in the primary lesions of CRC [34, 35]. The segmentation of primary lesions is greatly affected by bowel movement. Little is known regarding the radiomics features of metastatic liver lesions and their relationship to the genomic characteristics of RAS and BRAF.

In this study, we aimed to characterize the radiomics and semantic features in a multi-institutional cohort of patients with RAS- (KRAS and NRAS) and BRAF-mutated CRLM prior to any treatment. This study proved that the application of radiomics features together with semantic features, using an ANN method, can improve the non-invasive assessment of the mutation of RAS and BRAF in CRLM.

Materials and methods

Patient

The entire cohort was acquired from the January 2014 to October 2019 records of the institutional picture archiving and communication system (PACS, Philips), which was used to identify patients who had confirmed CRLM histologically. A total of 159 patients was confirmed to the criteria. The following are the inclusion criteria for this study: (1) the patient is older than 18 years; (2) the patient was confirmed as colorectal adenocarcinoma with hepatic metastasis by histopathological examination; (3) tissue samples of all primary tumors for the mutation analysis were acquired either through biopsy or surgical resection; (4) RAS and BRAF gene status were obtained by sequencing or hybridization methods; (5) abdominal contrast-enhanced CT (CECT) images before chemical or radical therapy were available; (6) the slice thickness of CT ≤ 2 mm (to avoid data inconsistency); (7) the intervals between lung and abdominal CT examination and histopathological diagnosis were less than 31 days (range 4-30 days). The following are the exclusion criteria for this study: (1) patients with more than one primary tumor location; (2) poor quality of CT images due to patients' breathing or movement artifacts; (3) patients with too blurred edges to delineate. Baseline demographic and clinical characteristics were collected from the PACS medical records. Clinical-pathological factors included age, sex, RAS (KRAS and NRAS) and BRAF status. RAS and BRAF genes are both in the RAS-RAF-MAPK pathway and previous studies had classified patients with any mutant RAS or BRAF genes into one category [36, 37]. Based on the status of RAS and BRAF genes, the patients were classified into two groups: the mutant group and the wild group. Patients with any mutant RAS or BRAF were classified into the mutant group (N

= 93), others were classified into the wild group (N = 66).

This retrospective study was performed in accordance with the principles of the Helsinki Declaration and was approved by the Ethics Committee of the First Hospital of China Medical University. All of the written informed consent was obtained from all enrolled patients in this study.

CT protocol

The contrast administration of abdomen CT scans are patient specific and based on clinical guidelines [38]. All CT examinations performing with 64-row spiral CT scanners were acquired according to standardized scanning protocols [39]. CT manufacturers included Elscint, GE, Phillips, Siemens, and Toshiba. A 1.2-1.5 mL/kg body weight bolus of iohexol was injected intravenously at a flow rate of 2.5 mL/s, followed by a 20-30 mL saline flush. Patients were imaged in the supine position at full inspiration [40, 41]. Portal venous phase (PVP) was obtained 60-70 s after intravenous injection of contrast [42, 43]. The scanning parameters were as follows: tube voltage was 120 kVp (range 100-140 kVp), slice thickness was 2 mm, matrix was 512 × 512, tube current was 333 mA (range 100-752 mA), exposure time 751 ms (range 500-1782 ms), a standard reconstruction algorithm.

All the steps are compliant to Image Biomarker Standardisation Initiative (IBSI) standard [44-47]. All CT images are stored in the DICOM format. Visual inspection of each images was conducted in advance to sending these images for radiomics analysis in order to ensure that collected images are qualified for being used for analysis [48].

Segmentation

Radiomics extracts high-throughput quantitative imaging features to perform subsequent data analysis related to target clinical outcome. The workflow of a typical radiomics process consists of four steps: tumor segmentation, feature extraction, score construction and score evaluation.

Regions of interest (ROIs) in the portal venous phase (PVP) CT images were segmented with a three-dimensional semi-automatic segmentation method by two radiologists. Radiomics features were extracted from within the defined ROI to quantify tumor intensity, shape and texture. Two semantic features ("micro-satellite" and the presence of metastatic lesions other than the liver and regional lymph nodes) were assessed by two radiologists. We constructed the semantic, radiomics and combined scores using seven machine learning algorithms. Three scores using ANN method showed the best predictive performance based on semantic, radiomics and combined features. The performances of the three constructed scores were assessed by the area under a receiver operating characteristic (ROC) curve followed by decision curve analysis.

PVP CT images were anonymized for all individual and institutional data and were labelled with a random number. Three-dimensional semi-automatic segmentation was performed by 2 radiologists with work experience of 9 years, using the open-source 3D-Slicer software (www.slicer.org). The metastatic liver lesions with the largest cross-sectional area and clear boundary were selected. First, the chosen liver lesions were contoured using soft tissue windows (window width: 350 HU, window level: 40 HU) using a semi-automatic fast-marching segmentation algorithm on all PVP images slice-by-slice for each patient, slightly along the visible borders of the lesion to include the entire lesion's volume approximated [49]. Then the radiologists manually modified slice-by-slice again to erase adjacent normal tissue or surrounding bile ducts. Thirdly, the final "ball" segmentation resulted as regions of interest (ROIs) were inspected by a senior radiologist with 17 years of work experience. CT images in the DICOM format were imported into the 3D-Slicer software, subsequently the ROIs were exported into NRRD and MRML format for storage and further analysis [50].

Radiomics features preprocessing and selection

Radiomics features of the ROIs were extracted using Pyradiomics package in 3D-Slicer software [51]. The extracted radiomics features could be divided into 3 kinds: first-order features, shape-based features and textural features. First-order features describe the distribu-

tion of voxel intensities within the ROIs. Shape-based features captured the direct-viewing characteristics of ROIs as two- and three-dimensional size and shape. These features are independent from the gray level intensity distribution in the ROIs. Textural features were extracted based on five textural matrixes: (1) Gray Level Co-occurrence Matrix (GLCM), (2) Gray Level Size Zone Matrix (GLSZM), (3) Gray Level Run Length Matrix (GLRLM), (4) Neighbouring Gray Tone Difference Matrix (NGTDM) and (5) Gray Level Dependence Matrix (GLDM). Besides, we used a wavelet filter onto the original image in order to suppress noise and to extract detailed, high-dimensional radiomics feature [52].

Features with an exact same value in all patients, which were not discriminative, were excluded [53]. In order to verify the robustness and stability of the radiomics features, the intraclass correlation coefficient (ICC) and concordance correlation coefficient (CCC) for every extracted radiomics features in 30 randomly selected ROIs were also calculated. To calculate ICC, 2 radiologists with work experience of 9 years performed the ROIs segmentation. Both of them knew about the diagnosis of CRLM but were blinded to the clinical or pathological information. To calculate CCC, one radiologist repeated the ROIs segmentation and radiomics features extraction twice in a two-week period. Features with an ICC or CCC lower than 0.75 were excluded for subsequent analysis.

Semantic features

Two representative semantic features were also included in this study. Most of the semantic features that have been reported in previous radomics-related studies on digestive tract cancers are the tumor size, whether the edges are clear, and whether there is a specific metastasis [54, 55]. First, the tumor size has been included in the radiomics features that we have analyzed. Secondly, previous literature suggested that lesions with too blurred edges would affect the accuracy of segmentation [49, 56]. Therefore, based on the exclusion criteria in the study, we excluded patients with too blurred edges to delineate to maximize the accuracy of the segmentation. Thirdly, several studies have documented that the treatment options and prognosis are different among liver metastases and other metastases in colorectal cancer [57-61]. Therefore, we summarized the first semantic feature as the presence of metastatic lesions other than the liver and regional lymph nodes.

Besides, the relationship between RAS/BRAF mutation and positive resection margins has been widely investigated [15, 62]. Therefore, we summarized the second semantic feature, "micro-satellite" phenomenon, referred that a single large lesion is surrounded by multiple small lesions. Two radiologists with over 5 years of working experience independently reviewed all pre-treatment CT images and assessed the presence or absence of two semantic features.

Score construction and evaluation

Three scores were constructed to predict gene mutation status: a semantic score, a radiomics score and a combined score. The combined score integrated both of the radiomics and semantic features. In the training cohort, seven machine learning methods were used to construct three scores predicting the gene mutation status, including Artificial Neural Network (ANN), Gaussian Naive Bayes (GNB), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression, AdaBoost, Gradient Boost Classifier. The prediction performance of the three scores afforded by several methods was illustrated analyzing the area under the receiver operating characteristic (ROC) curve (AUC) in primary and validation datasets using "pROC" package in R. Delong validation was used to compare AUCs in different scores and determine whether they differed significantly. Besides, accuracy, sensitivity and specificity were calculated among different algorithms in primary and validation datasets. Decision curve analysis was performed to illustrate decision benefit using "Decision Curve" package in R. Among these seven algorithms, the algorithm with the best prediction performance was selected.

Subgroup analysis

The subgroup analysis was conducted to assess whether the score had a better predictive performance among a particular type of patient [63]. The evaluation of the algorithm with the best prediction performance in several

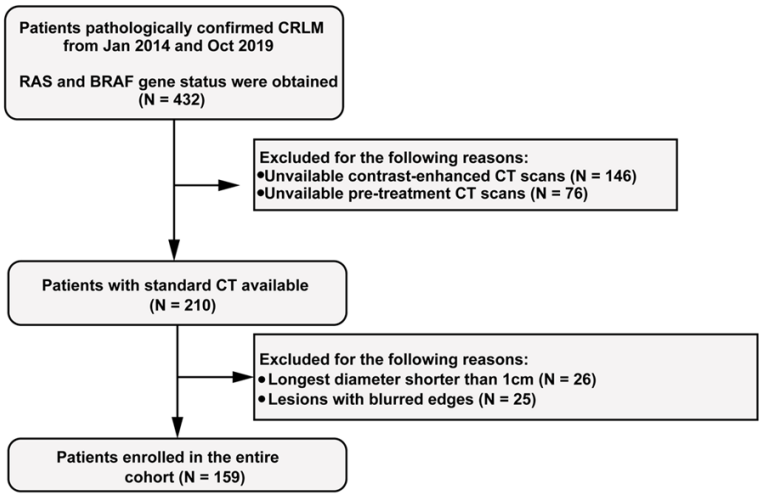

Figure 1. Flow chart of the enrolled patients in the study.

subgroups was conducted. Considering that RAS and BRAF mutations are not identical, we performed a subgroup analysis of patients with different mutant statuses in RAS and BRAF genes.

Analysis and comparison between different institutions

Since the entire cohort was acquired in multiple institutions, there were some differences which might directly affect the analysis of radiomics features between the two institutions. Therefore, we compared the prediction performance among patients from the two institutions to evaluate the generality of the three scores.

Statistical analysis

We used Pytorch (version 3.6.10) and R (version 3.5.3) for all statistical analysis. The Fisher's exact test was used to determine whether the clinical-pathological and semantic variables differed significantly between the primary and validation sets. Two-sided p values < 0.05 were considered statistically significant. Source code of statistical analysis can be accessed at: https://github.com/WissingChen/DeepRadiomics.

Results

Patient characteristics

One hundred and forty-eight patients (92 men and 56 women, median age of 60 years, age ranging between 35-79 years) from the First Hospital of China Medical University and 11 patients (5 men and 6 women, median age of 56 years, age ranging between 42-75 years) from the First Affiliated Hospital of Jinzhou Medical University were recruited to the study. **Figure 1** shows the patient recruitment process in this study. Of the 159 total patients, 124 patients were randomly allocated to the primary cohort, and 35 patients allocated to the validation cohort. The characteristics of the primary and validation cohorts are summarized in **Table 1**. No significant differences in age, sex, mutation status of KRAS, NRAS and BRAF genes, micro-satellite or extrahepatic metastasis were detected between the primary and validation cohorts (P = 0.12-0.99). No significant differences between the wild-type and mutant groups were found in the primary and validation datasets. Our study flow diagram is shown in **Figure 2**.

Construction and validation of the three scores

Of the 888 extracted radiomics features, 37 features with ICC and CCC lower than 0.75 were excluded from the subsequent analysis. Eight hundred and fifty-one features were retained for subsequent analysis. These radiomics features were classified into four categories: first-order features (N = 18), textural features (N = 75), shape-based features (N = 14) and wavelet features (N = 744).

Amongst the seven machine learning methods, three scores (semantic, radiomics and combined scores) using the ANN method showed the best predictive performance based on semantic, radiomics and combined features (**Figures 3**, S1 and Table S1). ANN fixed parameters of batch normalization and dropout layer, to reduce the possibility of overfitting and to guarantee the robustness of the radiomics features.

To evaluate the predictive performance, we applied the three scores based on the ANN algorithm in the primary and validation datasets. The AUC value of the radiomics score

Table 1. Characteristics of patients in the primary and validation cohorts

Characteristics	Primary Cohort (N = 124)	Validation Cohort (N = 35)	*P* Value
Age			> .99
≤ 55	45 (36.29%)	12 (34.29%)	
> 55	79 (63.71%)	23 (65.71%)	
Sex			0.85
Woman	49 (39.52%)	13 (37.14%)	
Man	75 (60.48%)	22 (62.86%)	
KRAS status			> .99
Mutant	64 (51.61%)	18 (51.43%)	
Wild	60 (48.39%)	17 (48.57%)	
NRAS status			0.12
Mutant	10 (8.06%)	0	
Wild	96 (77.42%)	30 (85.71%)	
NA	18 (14.52%)	5 (14.29%)	
BRAF status			0.69
Mutant	7 (5.65%)	1 (2.86%)	
Wild	110 (88.71%)	32 (91.43%)	
NA	7 (5.65%)	2 (5.71%)	
Microsatellite			0.15
No	83 (66.94%)	28 (80%)	
Yes	41 (33.06%)	7 (20%)	
Extrahepatic Meta.			0.84
No	82 (66.13%)	24 (68.57%)	
Yes	42 (33.87%)	11 (31.43%)	
Gene Status			0.57
Mutant	74 (59.68%)	19 (54.29%)	
Wild	50 (40.32%)	16 (45.71%)	

Note: The *p* values derived from the Fisher's exact test. "Micro-satellite" refers that a single large lesion is surrounded by multiple small lesions. "Extra-hepatic Meta". refers to presence of metastatic lesions other than the liver and regional lymph nodes.

reached 0.90 in the training dataset. After combining the radiomics with the effective semantic score, the predictive performance of the combined score was shown to be significantly improved with an AUC of 0.95 in the primary dataset and 0.79 in the validation dataset, showing the best discriminative efficacy (**Figure 4A** and **4B**).

The best-performed combined score and both of the radiomics and semantic scores were compared. The DeLong's test manifested a significant difference in the primary and validation datasets with *p* values < 0.05, separately. The accuracies of the combined scores in the primary and validation cohorts were 0.87 and 0.71, respectively. Decision curve analysis indicated that the combined score could improve the benefit compared with the measures that treat all patients and treat no patients (**Figure 4C** and **4D**).

Subgroup analysis

The performance of the combined score in several subgroups was displayed in **Table 2**. The accuracy of the combined score was relatively high in these subgroups, especially in the patients with mutant status of RAS and BRAF genes.

Analysis and comparison between different institutions

The performances of the combined score constructed using the ANN method were also evaluated amongst patients in two hospitals. The predictive performances of the combined scores, in terms of AUC, accuracy, sensitivity and specificity, were similar in cohorts of patients from both hospitals (Figure S2).

Discussion

Our analysis indicated that the lesions of RAS and BRAF mutant liver metastases exhibit radiomics and semantic features that allow them to be distinguished from wild-type tumors. In this study, we mined and validated a combined score that tracks RAS (KRAS and NRAS) and BRAF mutant phenotypes in CRLM from CT image data.

The steps used in this study attempt to strictly follow the radiomics quality score (RQS) proposed by Lambin and colleagues [64]. We provided well-documented CT protocols, analysis of feature robustness, combination of semantic features, analysis of discrimination and calibration, and a comparison to the current gold standard. The process of semi-automatic segmentation with 3D slicer was proven as a better substitution for manual segmentation during

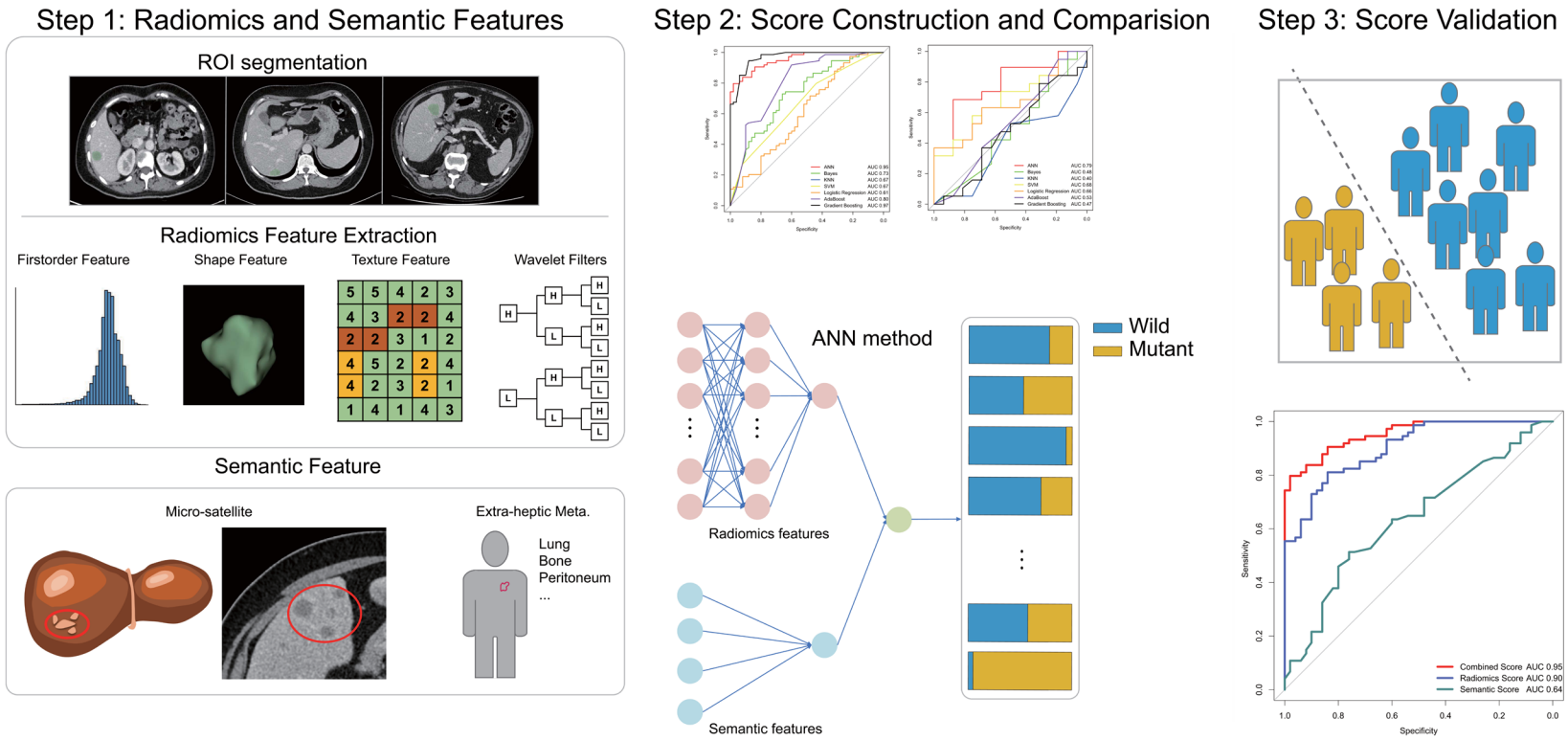

Figure 2. Workflow of the necessary steps in this study. Regions of interest (ROIs) in the portal venous phase (PVP). CT images were segmented with a three-dimensional, semi-automatic segmentation method by two radiologists. Radiomics features were extracted from the defined ROI to quantify tumor intensity, shape and texture. Two semantic features ("micro-satellite" and the presence of metastatic lesions other than the liver and regional lymph nodes) were assessed by two radiologists. We constructed the semantic, radiomics and combined scores using seven machine learning algorithms. Three scores using the artificial neural network (ANN) method showed the best predictive performance based on semantic, radiomics and combined features. The performances of the three constructed scores were evaluated by the area under a receiver operating characteristic (ROC) curves followed by decision curve analysis.

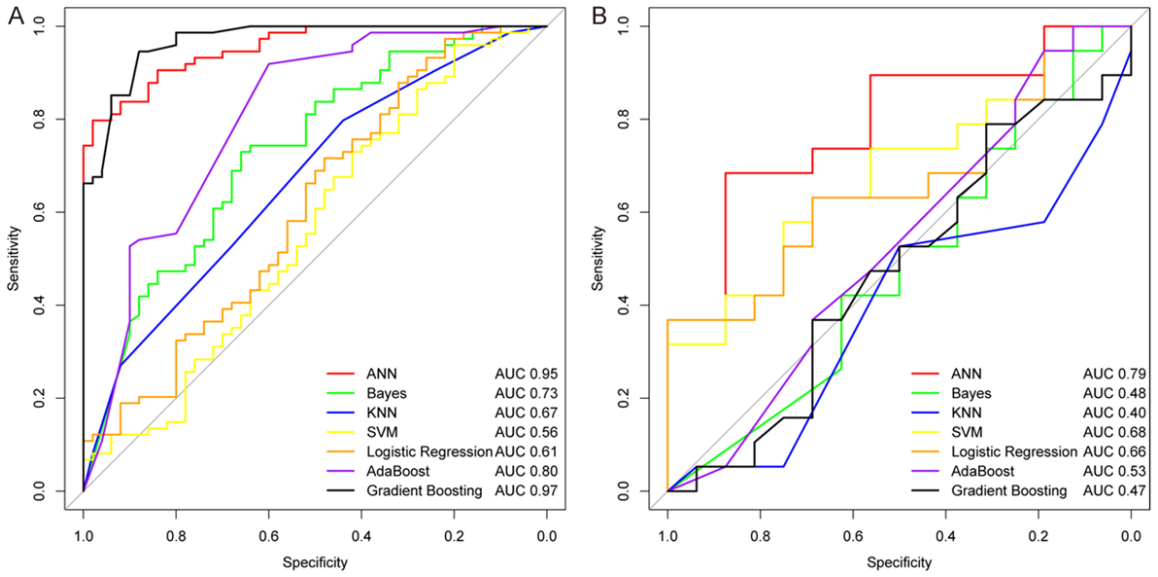

Figure 3. The ROC curves of the combined scores using seven methods in primary (A) and validation cohorts (B).

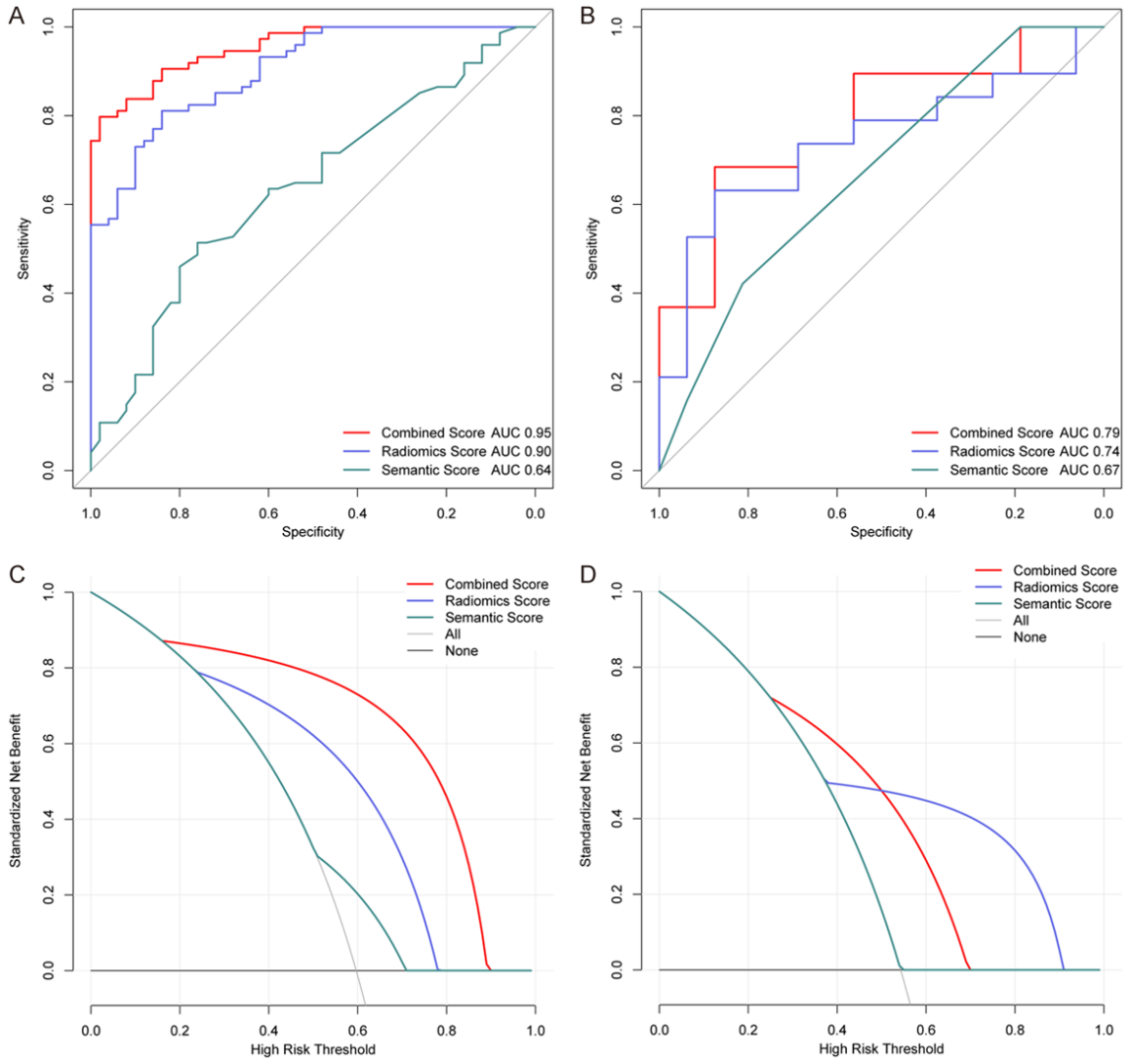

Figure 4. The ROC curves of the semantic, radiomics and combined scores using the ANN method in primary (A) and validation cohorts (B). Decision curve analysis for the three scores in the primary (C) and validation cohorts (D). The y-axis represents the net benefit and the x-axis denotes the threshold probability.

Table 2. Performance of the combined score using the ANN method

Subset	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Training set	87.10	89.19	84.00	89.19	84.00
Validation set	71.43	69.57	75.00	84.21	56.25
Micro-satellite	89.58	91.18	85.71	93.94	80.00
No micro-satellite	81.08	80.95	81.25	85.00	76.47
Extra-heptic Meta.	86.79	92.11	73.33	89.74	78.57
No Extra-heptic Meta.	82.08	79.66	85.11	87.04	76.92
RAS Wild-type	78.57	21.05	100	100	77.27
RAS Mutant-type	87.50	87.50	*	100	*
BRAF Wild-type	83.10	81.71	85.00	88.16	77.27
BRAF Mutant-type	100.00	100.00	*	100	*

Note: ANN: artificial neural network; PPV: positive predictive value; NPV: negative predictive value. *According to the combined score, all patients were predicted to be mutant patients.

quantification of radiomics features. This method of quantification could reduce uncertainty associated with tumor delineation and allowed more robust quantification of radiomics features with reduced processing time [39, 65]. In recent studies, 3D image features have been shown to provide equivalent predictive performance and better quantification of tumor heterogeneity compared to 2D image features [66-68]. Of seven different machine learning methods, we selected a multi-layer perception, which constructed three scores with the best predictive performance. Multi-layer perception is a deep Artificial Neural Network (ANN) with modern techniques such as weight decay regularization, Leaky Rectified Linear Units (ReLU) with batch normalization, dropout layer, Adaptive Moment Estimation (Adam), gradient clipping and learning rate scheduling. In the training cohort, we used a validation mode in PyTorch with L2 regularization to reduce the possibility of overfitting, and to guarantee the robustness of the radiomics features. Due to the fast convergence speed of ANN training and the large number of model parameters, it was potentially easy to overfit the data. The parameters with the best effect of validation set were selected as the saved model during the training process. The proposed combined score incorporating the radiomics and semantic features exhibited favorable discrimination in both the primary and the validation cohorts which strongly supports the high reliability and repeatability of the data in supporting the conclusions from this study.

For patients with liver metastasis which may be resected after chemotherapy, the net benefit of prior chemotherapy versus prior resection remains uncertain, and these patients need to be treated individually. When balancing remaining liver function and R0 resection, it is also unclear whether patients with CRLM harboring RAS and/or BRAF mutations should undergo chemotherapy to minimize the resection field and increase the resection margin whilst ensuring acceptable remaining liver tissue function.

In clinical practice, abdominal CECT is a routine method of evaluation for patients with CRLM. Thus, the combined score can be applied in almost any clinical setting where abdominal CECT is performed. Given these data, we confirmed the potential role of a non-invasive radiomics based method in aiding the subtle genomic evaluation in CRLM, especially for patients unfit for surgery or biopsy. According to National Comprehensive Cancer Network (NCCN) guidelines, the evaluation of RAS and BRAF mutations can be conducted on primary colorectal cancers or metastatic sites [69]. Previous studies have observed a high concordance value (ranging between 96-100%) of RAS and BRAF mutations between primary colorectal tumors and corresponding liver metastases, as well as in multiple liver metas-

tases [70-72]. The concordance of gene status was lower between primary lesion and lung metastases [73]. In the study, we selected a single liver metastasis as the ROI for analysis instead of the entire lesion, the primary CRC lesion or lung metastases. This approach could avoid deviation or poor repeatability caused by bowel movements and ensured the high consistency of gene status.

Despite the advantages of using a validation cohort, our study had some limitations. Firstly, the data in this study were retrospectively collected. Secondly, the ANN algorithm itself limited the interpretability of the results. Prospective investigation using considerably larger datasets and more detailed subgroup analysis is required to further validate the robustness and reproducibility of our conclusions. Despite these limitations, we believe the findings are robust and scalable to larger patient populations.

In conclusion, RAS (KRAS and NRAS) and BRAF mutated tumors show discriminative CT radiomics features that, when combined with semantic features, can aid the prediction of tumors harboring RAS and BRAF mutations to ultimately improve patient stratification and individualized treatments. Furthermore, we aim to successfully translate this understanding into clinical practice by allowing oncologists and surgeons to gain critical information regarding the molecular subtype of CRC, that can be adopted to directly inform better decision making in the clinic. These results have promising implications, radiomics and semantic features will require further validation in larger patient cohorts.

Acknowledgements

This work was supported by grants from the National Science and Technology Major Project of the Ministry of Science and Technology of China (No. 2017ZX09304025); Technological Special Project of Liaoning Province of China (2019020176-JH1/103); Science and Technology Plan Project of Liaoning Province (NO. 20-13225585); the General Projects of Liaoning Province Colleges and Universities (LFWK20-1706). The authors thank Yi Liu, Xiaoyi Hao, Zhicheng Chen for the research guidance. The authors also appreciate the experimental assistance given by Xi Chen, Yujia Song, Fang Wang, Liqing Jiang, Yujing Yang, Shiying Tang, Lingzi He, Hui Qu and Yunhui Wang.

Disclosure of conflict of interest

None.

Abbreviations

CRLM, colorectal cancer with liver metastasis; ROI, region of interest; CRC, colorectal cancer.

Address correspondence to: Xiujuan Qu and Zhi Li, Department of Medical Oncology, The First Hospital of China Medical University, 155 Nanjing Street, 110001, Liaoning Province, China. Tel: +86-13604031355; E-mail: xiujuanqu@yahoo.com (XJQ); Tel: +86-15840247601; E-mail: zli@cmu.edu.cn (ZL)

References

[1] Sun W, Ren S, Li R, Zhang Q and Song H. LncRNA, a novel target biomolecule, is involved in the progression of colorectal cancer. Am J Cancer Res 2019; 9: 2515-2530.

[2] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020; 70: 7-30.

[3] Jemal A, Ward EM, Johnson CJ, Cronin KA, Ma J, Ryerson B, Mariotto A, Lake AJ, Wilson R, Sherman RL, Anderson RN, Henley SJ, Kohler BA, Penberthy L, Feuer EJ and Weir HK. Annual report to the nation on the status of cancer, 1975-2014, featuring survival. J Natl Cancer Inst 2017; 109: djx030.

[4] Chen LY, Zhi Z, Wang L, Zhao YY, Deng M, Liu YH, Qin Y, Tian MM, Liu Y, Shen T, Sun LN and Li JM. NSD2 circular RNA promotes metastasis of colorectal cancer by targeting miR-199b-5p-mediated DDR1 and JAG1 signalling. J Pathol 2019; 248: 103-115.

[5] Joung JG, Oh BY, Hong HK, Al-Khalidi H, Al-Alem F, Lee HO, Bae JS, Kim J, Cha HU, Alotaibi M, Cho YB, Hassanain M, Park WY and Lee WY. Tumor heterogeneity predicts metastatic potential in colorectal cancer. Clin Cancer Res 2017; 23: 7209-7216.

[6] Camidge DR, Doebele RC and Kerr KM. Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. Nat Rev Clin Oncol 2019; 16: 341-355.

[7] Pan L and Cai X. PD-(L)1 inhibitors vs targeted therapy vs their combination as second-line treatment in patients with advanced hepatocellular carcinoma previously treated with sorafenib. Ann Oncol 2019; 30 Suppl 4: iv105.

[8] Li Q, Guan X, Chen S, Yi Z, Lan B, Xing P, Fan Y, Wang J, Luo Y, Yuan P, Cai R, Zhang P, Li Q, Zhong D, Zhang Y, Zou J, Zhu X, Ma F and Xu B.

Safety, efficacy, and biomarker analysis of pyrotinib in combination with capecitabine in HER2-positive metastatic breast cancer patients: a phase I clinical trial. Clin Cancer Res 2019; 25: 5212-5220.

[9] Loupakis F, Cremolini C, Salvatore L, Masi G, Sensi E, Schirripa M, Michelucci A, Pfanner E, Brunetti I, Lupi C, Antoniotti C, Bergamo F, Lonardi S, Zagonel V, Simi P, Fontanini G and Falcone A. FOLFOXIRI plus bevacizumab as first-line treatment in BRAF mutant metastatic colorectal cancer. Eur J Cancer 2014; 50: 57-63.

[10] Venook AP, Ou FS, Lenz HJ, Kabbarah O, Qu X, Niedzwiecki D, Zemla T, Goldberg RM, Hochster HS, O'Neil BH, Sanoff HK, Mayer RJ, Bertagnolli MM, Blanke CD and Innocenti F. Primary (1°) tumor location as an independent prognostic marker from molecular features for overall survival (OS) in patients (pts) with metastatic colorectal cancer (mCRC): analysis of CALGB/SWOG 80405 (Alliance). J Clin Oncol 2017; 35: 3503-3503.

[11] Modest DP, Stintzing S, Weikersthal LFv, Decker T, Kiani A, Vehling-Kaiser U, Al-Batran SE, Heintges T, Kahl C, Seipelt G, Kullmann F, Stauch M, Scheithauer W, Held S, Giessen-Jung C, Jung A, Kirchner T and Heinemann V. Primary tumor location and efficacy of second-line therapy after initial treatment with FOLFIRI in combination with cetuximab or bevacizumab in patients with metastatic colorectal cancer- FIRE-3 (AIOKRK0306). J Clin Oncol 2017; 35: 3525-3525.

[12] Bruera G, Cannita K, Di Giacomo D, Lamy A, Troncone G, Dal Mas A, Coletti G, Frebourg T, Sabourin JC, Tosi M, Ficorella C and Ricevuto E. Prognostic value of KRAS genotype in metastatic colorectal cancer (MCRC) patients treated with intensive triplet chemotherapy plus bevacizumab (FIr-B/FOx) according to extension of metastatic disease. BMC Med 2012; 10: 135.

[13] Cady B, Jenkins RL, Steele GD Jr, Lewis WD, Stone MD, McDermott WV, Jessup JM, Bothe A, Lalor P, Lovett EJ, Lavin P and Linehan DC. Surgical margin in hepatic resection for colorectal metastasis: a critical and improvable determinant of outcome. Ann Surg 1998; 227: 566-571.

[14] Ekberg H, Tranberg KG, Andersson R, Lundstedt C, Hägerstrand I, Ranstam J and Bengmark S. Determinants of survival in liver resection for colorectal secondaries. Br J Surg 1986; 73: 727-731.

[15] Brudvik KW, Mise Y, Chung MH, Chun YS, Kopetz SE, Passot G, Conrad C, Maru DM, Aloia TA and Vauthey JN. RAS mutation predicts positive resection margins and narrower resection margins in patients undergoing resection of colorectal liver metastases. Ann Surg Oncol 2016; 23: 2635-2643.

[16] Xu D, Wang HW, Yan XL, Li J, Wang K and Xing BC. Sub-millimeter surgical margin is acceptable in patients with good tumor biology after liver resection for colorectal liver metastases. Eur J Surg Oncol 2019; 45: 1551-1558.

[17] Margonis GA, Buettner S, Andreatos N, Sasaki K, Ijzermans JNM, van Vugt JLA, Pawlik TM, Choti MA, Cameron JL, He J, Wolfgang CL and Weiss MJ. Anatomical resections improve disease-free survival in patients with KRAS-mutated colorectal liver metastases. Ann Surg 2017; 266: 641-649.

[18] Okuno M, Goumard C, Kopetz S, Vega EA, Joechle K, Mizuno T, Omichi K, Tzeng CD, Chun YS, Vauthey JN and Conrad C. RAS mutation is associated with unsalvageable recurrence following hepatectomy for colorectal cancer liver metastases. Ann Surg Oncol 2018; 25: 2457-2466.

[19] Margonis GA, Buettner S, Andreatos N, Kim Y, Wagner D, Sasaki K, Beer A, Schwarz C, Løes IM, Smolle M, Kamphues C, He J, Pawlik TM, Kaczirek K, Poultsides G, Lønning PE, Cameron JL, Burkhart RA, Gerger A, Aucejo FN, Kreis ME, Wolfgang CL and Weiss MJ. Association of BRAF mutations with survival and recurrence in surgically treated patients with metastatic colorectal liver cancer. JAMA Surg 2018; 153: e180996.

[20] Douillard JY, Oliner KS, Siena S, Tabernero J, Burkes R, Barugel M, Humblet Y, Bodoky G, Cunningham D, Jassem J, Rivera F, Kocákova I, Ruff P, Błasińska-Morawiec M, Šmakal M, Canon JL, Rother M, Williams R, Rong A, Wiezorek J, Sidhu R and Patterson SD. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. N Engl J Med 2013; 369: 1023-1034.

[21] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A and Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48: 441-446.

[22] Liu H, Zhang C, Wang L, Luo R, Li J, Zheng H, Yin Q, Zhang Z, Duan S, Li X and Wang D. MRI radiomics analysis for predicting preoperative synchronous distant metastasis in patients with rectal cancer. Eur Radiol 2019; 29: 4418-4426.

[23] Forghani R, Chatterjee A, Reinhold C, Perez-Lara A, Romero-Sanchez G, Ueno Y, Bayat M, Alexander JWM, Kadi L, Chankowsky J, Seuntjens J and Forghani B. Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT

texture analysis with machine learning. Eur Radiol 2019; 29: 6172-6181.
[24] Hyun SH, Ahn MS, Koh YW and Lee SJ. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. Clin Nucl Med 2019; 44: 956-960.
[25] Zhang W, Fang M, Dong D, Wang X, Ke X, Zhang L, Hu C, Guo L, Guan X, Zhou J, Shan X and Tian J. Development and validation of a CT-based radiomic nomogram for preoperative prediction of early recurrence in advanced gastric cancer. Radiother Oncol 2019; 145: 13-20.
[26] Wang XH, Long LH, Cui Y, Jia AY, Zhu XG, Wang HZ, Wang Z, Zhan CM, Wang ZH and Wang WH. MRI-based radiomics model for preoperative prediction of 5-year survival in patients with hepatocellular carcinoma. Br J Cancer 2020; 122: 978-985.
[27] Antunovic L, De Sanctis R, Cozzi L, Kirienko M, Sagona A, Torrisi R, Tinterri C, Santoro A, Chiti A, Zelic R and Sollini M. PET/CT radiomics in breast cancer: promising tool for prediction of pathological response to neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging 2019; 46: 1468-1477.
[28] Bhatia A, Birger M, Veeraraghavan H, Um H, Tixier F, McKenney AS, Cugliari M, Caviasco A, Bialczak A, Malani R, Flynn J, Zhang Z, Yang TJ, Santomasso BD, Shoushtari AN and Young RJ. MRI radiomic features are associated with survival in melanoma brain metastases treated with immune checkpoint inhibitors. Neuro Oncol 2019; 21: 1578-1586.
[29] Park JE, Kim HS, Park SY, Nam SJ, Chun SM, Jo Y and Kim JH. Prediction of core signaling pathway by using diffusion- and perfusion-based MRI radiomics and next-generation sequencing in isocitrate dehydrogenase wild-type glioblastoma. Radiology 2020; 294: 388-397.
[30] Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, Leijenaar RT, Haibe-Kains B, Lambin P, Gillies RJ and Aerts HJ. Defining the biological basis of radiomic phenotypes in lung cancer. Elife 2017; 6: e23421.
[31] Li Y, Eresen A, Shangguan J, Yang J, Lu Y, Chen D, Wang J, Velichko Y, Yaghmai V and Zhang Z. Establishment of a new non-invasive imaging prediction model for liver metastasis in colon cancer. Am J Cancer Res 2019; 9: 2482-2492.
[32] Li Y, Eresen A, Lu Y, Yang J, Shangguan J, Velichko Y, Yaghmai V and Zhang Z. Radiomics signature for the preoperative assessment of stage in advanced colon cancer. Am J Cancer Res 2019; 9: 1429-1438.
[33] Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.
[34] Kim SJ, Pak K and Kim K. Diagnostic performance of F-18 FDG PET/CT for prediction of KRAS mutation in colorectal cancer patients: a systematic review and meta-analysis. Abdom Radiol (NY) 2019; 44: 1703-1711.
[35] Lovinfosse P, Koopmansch B, Lambert F, Jodogne S, Kustermans G, Hatt M, Visvikis D, Seidel L, Polus M, Albert A, Delvenne P and Hustinx R. (18)F-FDG PET/CT imaging in rectal cancer: relationship with the RAS mutational status. Br J Radiol 2016; 89: 20160212.
[36] van Helden EJ, Angus L, Menke-van der Houven van Oordt CW, Heideman DAM, Boon E, van Es SC, Radema SA, van Herpen CML, de Groot DJA, de Vries EGE, Jansen MPHM, Sleijfer S and Verheul HMW. RAS and BRAF mutations in cell-free DNA are predictive for outcome of cetuximab monotherapy in patients with tissue-tested RAS wild-type advanced colorectal cancer. Mol Oncol 2019; 13: 2361-2374.
[37] Modest DP, Ricard I, Heinemann V, Hegewisch-Becker S, Schmiegel W, Porschen R, Stintzing S, Graeven U, Arnold D, von Weikersthal LF, Giessen-Jung C, Stahler A, Schmoll HJ, Jung A, Kirchner T, Tannapfel A and Reinacher-Schick A. Outcome according to KRAS-, NRAS- and BRAF-mutation as well as KRAS mutation variants: pooled analysis of five randomized trials in metastatic colorectal cancer by the AIO colorectal cancer study group. Ann Oncol 2016; 27: 1746-1753.
[38] Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, Mak RH and Aerts H. Deep learning predicts lung cancer treatment response from serial medical imaging. Clin Cancer Res 2019; 25: 3266-3275.
[39] Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, Quackenbush J and Aerts H. Somatic mutations drive distinct imaging phenotypes in lung cancer. Cancer Res 2017; 77: 3922-3930.
[40] Xu Y, Lu L, E LN, Lian W, Yang H, Schwartz LH, Yang ZH and Zhao B. Application of radiomics in predicting the malignancy of pulmonary nodules in different sizes. AJR Am J Roentgenol 2019; 213: 1213-1220.
[41] Bashir U, Kawa B, Siddique M, Mak SM, Nair A, McLean E, Bille A, Goh V and Cook G. Non-invasive classification of non-small cell lung cancer: a comparison between random forest models utilising radiomic and semantic features. Br J Radiol 2019; 92: 20190159.
[42] Trebeschi S, Drago SG, Birkbak NJ, Kurilova I, Calin AM, Delli Pizzi A, Lalezari F, Lambregts DMJ, Rohaan MW, Parmar C, Rozeman EA, Hartemink KJ, Swanton C, Haanen JBAG, Blank CU, Smit EF, Beets-Tan RGH and Aerts HJWL. Predicting response to cancer immunotherapy

using non-invasive radiomic biomarkers. Ann Oncol 2019; 30: 998-1004.
[43] Sandrasegaran K, Lin Y, Asare-Sawiri M, Taiyini T and Tann M. CT texture analysis of pancreatic cancer. Eur Radiol 2019; 29: 1067-1073.
[44] Kirienko M, Cozzi L, Rossi A, Voulaz E, Antunovic L, Fogliata A, Chiti A and Sollini M. Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions. Eur J Nucl Med Mol Imaging 2018; 45: 1649-1660.
[45] Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ and Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014; 5: 4006.
[46] Vallieres M, Zwanenburg A, Badic B, Cheze Le Rest C, Visvikis D and Hatt M. Responsible radiomics research for faster clinical translation. J Nucl Med 2018; 59: 189-193.
[47] Zwanenburg A, Leger S, Vallières M and Löck S. Image biomarker standardisation initiative. Radiotherapy and Oncology 2016.
[48] Sasaki T, Kinoshita M, Fujita K, Fukai J, Hayashi N, Uematsu Y, Okita Y, Nonaka M, Moriuchi S, Uda T, Tsuyuguchi N, Arita H, Mori K, Ishibashi K, Takano K, Nishida N, Shofuda T, Yoshioka E, Kanematsu D, Kodama Y, Mano M, Nakao N and Kanemura Y. Radiomics and MGMT promoter methylation for prognostication of newly diagnosed glioblastoma. Sci Rep 2019; 9: 14435.
[49] Ma X, Wei J, Gu D, Zhu Y, Feng B, Liang M, Wang S, Zhao X and Tian J. Preoperative radiomics nomogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced CT. Eur Radiol 2019; 29: 3595-3605.
[50] Zhao W, Yang J, Ni B, Bi D, Sun Y, Xu M, Zhu X, Li C, Jin L, Gao P, Wang P, Hua Y and Li M. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. Cancer Med 2019; 8: 3532-3543.
[51] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S and Aerts HJWL. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017; 77: e104-e107.
[52] Cselényi Z, Olsson H, Farde L and Gulyás B. Wavelet-aided parametric mapping of cerebral dopamine D2 receptors using the high affinity PET radioligand [11C]FLB 457. Neuroimage 2002; 17: 47-60.
[53] de Jong EEC, Sanders KJC, Deist TM, van Elmpt W, Jochems A, van Timmeren JE, Leijenaar RTH, Degens JHRJ, Schols AMWJ, Dingemans AC and Lambin P. Can radiomics help to predict skeletal muscle response to chemotherapy in stage IV non-small cell lung cancer? Eur J Cancer 2019; 120: 107-113.
[54] Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, Zhang YD, Jiang WJ, Li XC and Wang XH. Biliary tract cancer at CT: a radiomics-based model to predict lymph node metastasis and survival outcomes. Radiology 2019; 290: 90-98.
[55] Jiang Y, Wang H, Wu J, Chen C, Yuan Q, Huang W, Li T, Xi S, Hu Y, Zhou Z, Xu Y, Li G and Li R. Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. Ann Oncol 2020; [Epub ahead of print].
[56] Ravanelli M, Agazzi GM, Tononcelli E, Roca E, Cabassa P, Baiocchi G, Berruti A, Maroldi R and Farina D. Texture features of colorectal liver metastases on pretreatment contrast-enhanced CT may predict response and prognosis in patients treated with bevacizumab-containing chemotherapy: a pilot study including comparison with standard chemotherapy. Radiol Med 2019; 124: 877-886.
[57] Tang W, Ren L, Liu T, Ye Q, Wei Y, He G, Lin Q, Wang X, Wang M, Liang F, Cui Y and Xu J. Bevacizumab plus mFOLFOX6 versus mFOLFOX6 alone as first-line treatment for RAS mutant unresectable colorectal liver-limited metastases: the BECOME randomized controlled trial. J Clin Oncol 2020; 38: 3175-3184.
[58] Goéré D, Glehen O, Quenet F, Guilloit JM, Bereder JM, Lorimier G, Thibaudeau E, Ghouti L, Pinto A, Tuech JJ, Kianmanesh R, Carretier M, Marchal F, Arvieux C, Brigand C, Meeus P, Rat P, Durand-Fontanier S, Mariani P, Lakkis Z, Loi V, Pirro N, Sabbagh C, Texier M and Elias D. Second-look surgery plus hyperthermic intraperitoneal chemotherapy versus surveillance in patients at high risk of developing colorectal peritoneal metastases (PROPHYLOCHIP-PRODIGE 15): a randomised, phase 3 study. Lancet Oncol 2020; 21: 1147-1154.
[59] Nozue M, Oshiro Y, Kurata M, Seino K, Koike N, Kawamoto T, Taniguchi H, Todoroki T and Fukao K. Treatment and prognosis in colorectal cancer patients with bone metastasis. Oncol Rep 2002; 9: 109-112.
[60] Ikoma N, You YN, Bednarski BK, Rodriguez-Bigas MA, Eng C, Das P, Kopetz S, Messick C, Skibber JM and Chang GJ. Impact of recurrence and salvage surgery on survival after multidisciplinary treatment of rectal cancer. J Clin Oncol 2017; 35: 2631-2638.

[61] Margalit O, Shacham-Shmueli E, Lawrence YR, Yang YX, Reiss KA, Golan T, Mamtani R, Halpern N, Aderka D, Giantonio B and Boursi B. Lung metastasis predicts better prognosis in metastatic colorectal cancer with mutated KRAS. Clin Colorectal Cancer 2019; 18: e300-e307.

[62] Gagnière J, Dupré A, Gholami SS, Pezet D, Boerner T, Gönen M, Kingham TP, Allen PJ, Balachandran VP, De Matteo RP, Drebin JA, Yaeger R, Kemeny NE, Jarnagin WR and D'Angelica MI. Is hepatectomy justified for BRAF mutant colorectal liver metastases? A multi-institutional analysis of 1497 patients. Ann Surg 2020; 271: 147-154.

[63] Pocock SJ and Stone GW. The primary outcome is positive - is that good enough? N Engl J Med 2016; 375: 971-979.

[64] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE and Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017; 14: 749-762.

[65] Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, Mitra S, Shankar BU, Kikinis R, Haibe-Kains B, Lambin P and Aerts HJ. Robust radiomics feature quantification using semiautomatic volumetric segmentation. PLoS One 2014; 9: e102107.

[66] Fave X, Cook M, Frederick A, Zhang L, Yang J, Fried D, Stingo F and Court L. Preliminary investigation into sources of uncertainty in quantitative imaging features. Comput Med Imaging Graph 2015; 44: 54-61.

[67] Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L and Schwartz LH. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep 2016; 6: 23428.

[68] Ng F, Kozarski R, Ganeshan B and Goh V. Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? Eur J Radiol 2013; 82: 342-348.

[69] Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen YJ, Ciombor KK, Cohen S, Cooper HS, Deming D, Engstrom PF, Garrido-Laguna I, Grem JL, Grothey A, Hochster HS, Hoffe S, Hunt S, Kamel A, Kirilcuk N, Krishnamurthi S, Messersmith WA, Meyerhardt J, Miller ED, Mulcahy MF, Murphy JD, Nurkin S, Saltz L, Sharma S, Shibata D, Skibber JM, Sofocleous CT, Stoffel EM, Stotsky-Himelfarb E, Willett CG, Wuthrick E, Gregory KM and Freedman-Cass DA. NCCN guidelines insights: colon cancer, version 2.2018. J Natl Compr Canc Netw 2018; 16: 359-369.

[70] Etienne-Grimaldi MC, Formento JL, Francoual M, François E, Formento P, Renée N, Laurent-Puig P, Chazal M, Benchimol D, Delpero JR, Letoublon C, Pezet D, Seitz JF and Milano G. K-Ras mutations and treatment outcome in colorectal cancer patients receiving exclusive fluoropyrimidine therapy. Clin Cancer Res 2008; 14: 4830-4835.

[71] Artale S, Sartore-Bianchi A, Veronese SM, Gambi V, Sarnataro CS, Gambacorta M, Lauricella C and Siena S. Mutations of KRAS and BRAF in primary and matched metastatic sites of colorectal cancer. J Clin Oncol 2008; 26: 4217-4219.

[72] Knijn N, Mekenkamp LJ, Klomp M, Vink-Börger ME, Tol J, Teerenstra S, Meijer JW, Tebar M, Riemersma S, van Krieken JH, Punt CJ and Nagtegaal ID. KRAS mutation analysis: a comparison between primary tumours and matched liver metastases in 305 colorectal cancer patients. Br J Cancer 2011; 104: 1020-1026.

[73] Moorcraft SY, Jones T, Walker BA, Ladas G, Kalaitzaki E, Yuan L, Begum R, Eltahir Z, Wotherspoon A, Montero-Fernandez A, Teixeira Mendes LS, Gonzalez de Castro D, Wilson SH, Proszek P, To YM, Hawkes E, Roy A, Cunningham D, Rao S, Watkins D, Starling N, Bowcock AM and Chau I. Molecular profiling of colorectal pulmonary metastases and primary tumours: implications for targeted treatment. Oncotarget 2017; 8: 64999-65008.

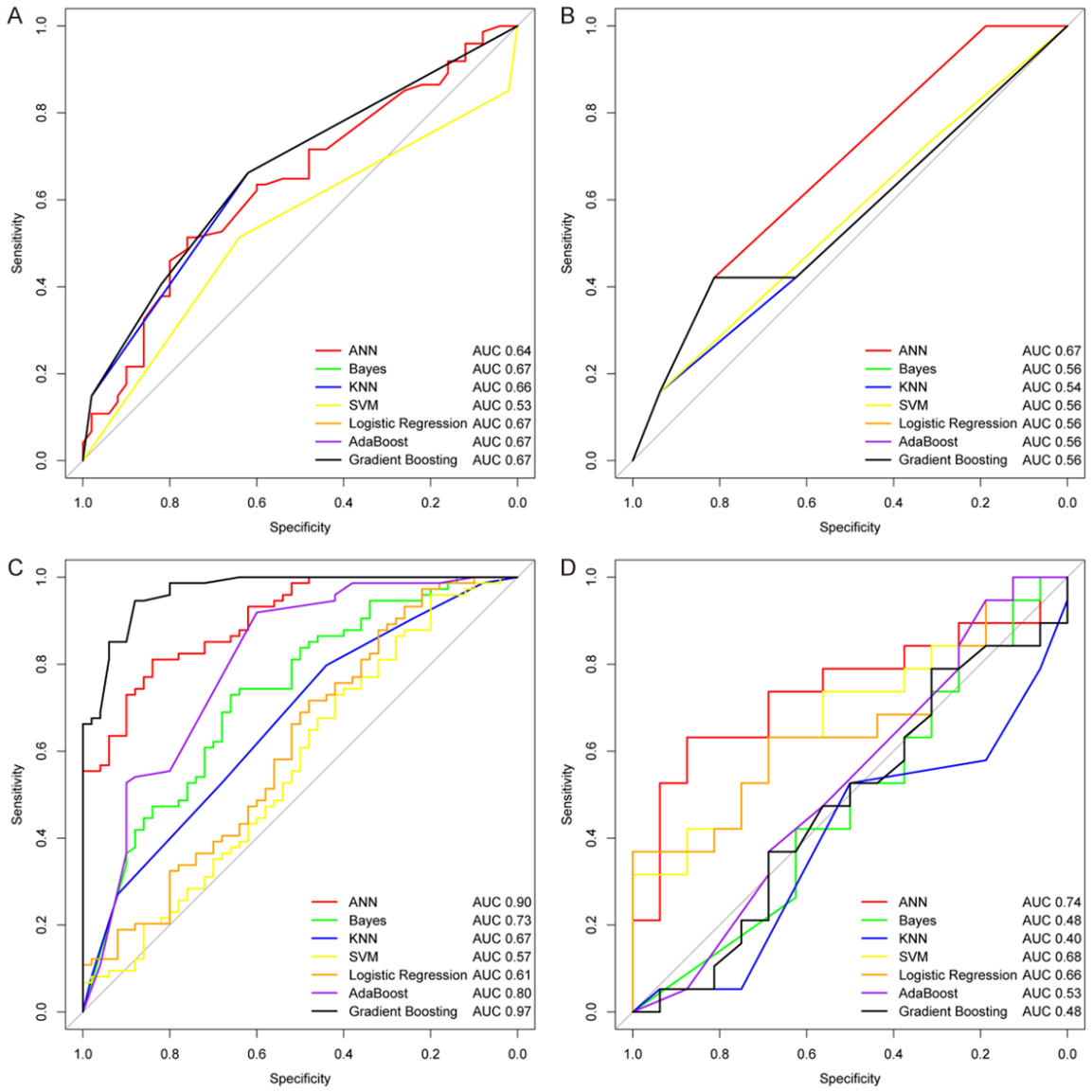

Figure S1. The ROC curves of the semantic (A, B) and radiomics scores (C, D) using seven methods in the primary (A, C) and validation cohorts (B, D). Note: ANN: artificial neural network; KNN: K-Nearest Neighbors; SVM: Support Vector Machine.

Table S1. Performance of combined score using different methods

Methods	Cohorts	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
ANN	Training cohort	87.10	89.19	84.00	89.19	84.00
	Validation cohort	71.43	69.57	75.00	84.21	56.25
Bayes	Training cohort	68.55	68.82	67.74	86.49	42.00
	Validation cohort	54.29	55.17	50.00	84.21	18.75
KNN	Training cohort	63.71	63.81	63.16	90.54	24.00
	Validation cohort	62.86	60.00	80.00	94.74	25.00
SVM	Training cohort	65.32	63.96	76.92	95.95	20.00
	Validation cohort	60.00	58.06	75.00	94.74	18.75
Logistic Regression	Training cohort	60.48	60.16	100	100	2.00
	Validation cohort	57.14	55.88	100	100	6.25
AdaBoost	Training cohort	73.39	70.71	84.00	94.59	42.00
	Validation cohort	60.00	58.06	75.00	94.74	18.75
Gradient Boosting	Training cohort	75.81	71.15	100	100	40.00
	Validation cohort	57.14	56.25	66.67	94.74	12.50

Note: PPV: positive predictive value; NPV: negative predictive value; ANN: artificial neural network; KNN: K-Nearest Neighbors; SVM: Support Vector Machine.

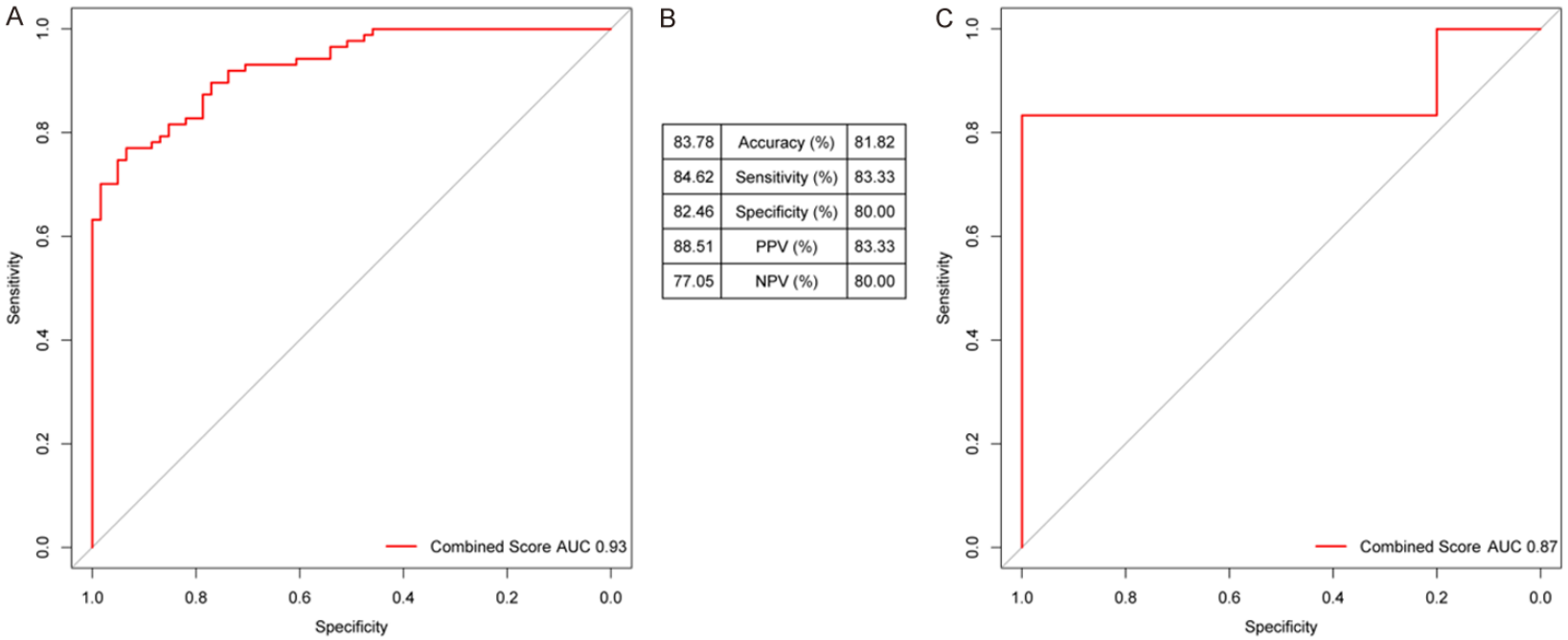

Figure S2. The predictive performance of the combined score amongst patients from different hospitals. The ROC curves of the combined score using the ANN method among patients from one hospital (A) and another hospital (C). (B) The accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the combined score. The right volume refers to one hospital and the left volume represents another hospital.