Heliyon 10 (2024) e32910

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

CT-based radiomics predicts CD38 expression and indirectly reflects clinical prognosis in epithelial ovarian cancer

Yuan Yao [a], Haijin Zhang [a], Hui Liu [a], Chendi Teng [b], Xuan Che [c,*], Wei Bian [a], Wenting Zhang [a], Zhifeng Wang [a]

[a] Department of Radiology, Jiaxing Maternity and Child Health Care Hospital, Jiaxing, Zhejiang 314000, China
[b] Department of Radiology, Wenzhou Central Hospital, Wenzhou, Zhejiang, 325000, China
[c] Department of Gynecology, Jiaxing Maternity and Child Health Care Hospital, Jiaxing, Zhejiang, 314000, China

ARTICLE INFO

Keywords:
CD38
Epithelial ovarian cancer
Radiomics
CT
Survival
Prognosis

ABSTRACT

Background: Cluster of differentiation 38 (CD38) has been found to be highly expressed in various solid tumours, and its expression level may be associated with patient prognosis and survival. This study aimed to evaluate the prognostic value of CD38 expression for patients with epithelial ovarian cancer (EOC) and construct two computed tomography (CT)-based radiomics models for predicting CD38 expression.

Methods: A total of 333 cases of EOC were enrolled from The Cancer Genome Atlas (TCGA) database for CD38-related bioinformatics and survival analysis. A total of 56 intersection cases from TCGA and The Cancer Imaging Archive (TCIA) databases were selected for radiomics feature extraction and model construction. Logistic regression (LR) and support vector machine (SVM) models were constructed and internally validated using 5-fold cross-validation to assess the performance of the models for CD38 expression levels.

Results: High CD38 expression was an independent protective factor (HR = 0.540) for overall survival (OS) in EOC patients. Five radiomics features based on CT images were selected to build models for the prediction of CD38 expression. In the training and internal validation sets, for the receiver operating characteristic (ROC) curve, the LR model reached an area under the curve (AUC) of 0.739 and 0.732, while the SVM model achieved AUC values of 0.741 and 0.700, respectively. For the precision-recall (PR) curve, the LR and SVM models demonstrated an AUC of 0.760 and 0.721. The calibration curves and decision curve analysis (DCA) provided evidence supporting the fitness and net benefit of the models.

Abbreviations: CD38, Cluster of differentiation 38; EOC, Epithelial ovarian cancer; CT, Computed tomography; TCGA, The Cancer Genome Atlas; TCIA, The Cancer Imaging Archive; LR, Logistic regression; SVM, Support vector machine; OS, Overall survival; ROC, Receiver operating characteristic; AUC, Area under the curve; PR Curve, Precision-recall curve; DCA, Decision curve analysis; CA 125, Carbohydrate antigen 125; HE 4, Human epididymis protein 4; MRI, Magnetic resonance imaging; DFS, Disease-free survival; mRNA, Message ribonucleic acid; RNA-seq, Ribonucleic acid sequencing; OC, Ovarian cancer; OSC, Ovarian serous cystadenocarcinoma; FIGO, International Federation of Gynecology and Obstetrics; GTEx database, Genotype-tissue expression database; GSEA, Gene set enrichment analysis; KEGG, The Kyoto Encyclopedia of Genes and Genomes; VOIs, The volumes of interest; ICC, Intraclass correlation coefficient; mRMR, Maximum relevance minimum redundancy; AIC, Akaike information criterion; Rad_score, Radiomics score; NK cell, Natural killer cell; DC, Dendritic cell; γδT cell, Gamma-delta T cell; MAPK, signalling pathway, Mitogen-activated protein kinase signalling pathway..

* Corresponding author.
E-mail address: chexuan@zjxu.edu.cn (X. Che).

https://doi.org/10.1016/j.heliyon.2024.e32910
Received 13 January 2024; Received in revised form 29 May 2024; Accepted 11 June 2024
Available online 13 June 2024

Conclusions: High levels of CD38 expression can improve OS in EOC patients. CT-based radiomics models can be a new predictive tool for CD38 expression, offering possibilities for individualised survival assessment for patients with EOC.

1. Background

Epithelial ovarian cancer (EOC) is the most common type of ovarian malignancy, accounting for over 95 % of cases [1]. Annually, approximately 230,000 women are diagnosed with EOC, with more than 75 % of patients being diagnosed at an advanced stage. Furthermore, this disease claims the lives of around 150,000 patients per year worldwide [2]. The standard treatment for EOC involves surgical intervention followed by platinum-based adjuvant chemotherapy. However, in recent years, advances in immunotherapy and targeted therapy have emerged as promising treatment options [3]. Despite these advances, the average 5-year survival rate for EOC patients remains below 50 %, especially for those diagnosed at an advanced stage, where the 5-year survival rate is a mere 29 % [1,2]. Moreover, the prognosis and survival outcomes vary significantly among individuals [2,4], attributed to the considerable heterogeneity observed both among patients and within tumours [1,2]. The current diagnostic and prognostic indicators of EOC, including clinical and pathological characteristics [5], serum markers, such as carbohydrate antigen 125 (CA125) and human epididymis protein 4 (HE4) [6], and traditional techniques, such as ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI), are inadequate to meet the clinical demands of precision medicine [7]. Consequently, there is a pressing need to explore novel prognostic indicators that can facilitate the early and effective assessment of patients and support clinical decision-making.

The cluster of differentiation 38 (CD38) protein encoded by the CD38 gene is a non-lineage-restricted type II transmembrane glycoprotein that synthesises and hydrolyses cyclic adenosine 5′-diphosphate ribose, serving as an intracellular calcium mobilisation regulator. CD38 is expressed on the surface of various immune cells, promoting immune cell activation, proliferation, and adhesion [8], as well as exerting certain regulatory effects on tumour development [9]. Studies have found that CD38 is also highly expressed in lymphoid neoplasm cells, especially multiple myeloma (MM) [10], and solid tumour cells, such as head and neck squamous cell carcinoma [11], cervical cancer [12], and prostate cancer [13]. Furthermore, it has been confirmed to be associated with immune cell infiltration and tumour metabolism regulation [9]. Furthermore, Zhu et al. [14] found that CD38 expression in EOC tissues was higher than in normal tissues, and high CD38 expression has been reported to be associated with longer disease-free survival (DFS) in EOC patients. The biological associations of CD38 with EOC were also investigated by several studies, revealing that increased CD38 expression is associated with the enrichment of antitumor immune gene signatures in various categories, including immune response, lymphocyte activation, regulation of T cell-mediated immunity, and NK cell-mediated cytotoxicity, among EOC patients [14]. Furthermore, heightened CD38 expression showed a positive correlation with the infiltration of tumour-infiltrating lymphocytes, such as $CD8^+$ T cells, $CD4^+$ T cells, and B cells, in the EOC microenvironment, indicative of a potential contribution to the modulation of antitumor immunity [15,16]. Therefore, CD38 shows potential as both a biomarker for EOC prognosis and as a therapeutic target [17,18]. However, the currently available methods for detecting the CD38 protein or messenger ribonucleic acid (mRNA), which include peripheral blood cytokine detection, flow cytometry, ribonucleic acid sequencing (RNA-seq), and pathological immunohistochemistry, although mature and accurate, have drawbacks that make them unsuitable for frequent clinical application. These include invasiveness, operational complexity, relatively high costs, and an inability to comprehensively reflect the expression profile of tumour tissues. Therefore, there is a need to develop non-invasive and efficient methods that are able to reflect molecular expression levels in the whole tumour.

Currently, CT is one of the most commonly used imaging methods for the diagnosis, staging, and follow-up of ovarian cancer (OC) [7]. CT images are easily accessible and provide valuable information on the morphology of lesions, their surrounding infiltration, and the extent of metastasis. However, traditional imaging methods, such as CT and MRI, rely on visual identification and have limitations in providing comprehensive information [19]. Radiomics, on the other hand, has the potential to overcome these limitations by converting conventional medical image information into high-dimensional data that can be analysed quantitatively. Radiomics provides a wide range of quantitative features based on intensity, shape, volume, and texture, which are closely related to molecular-level characteristics and the tumour microenvironment [19]. This methodology has been successfully applied in various disease states for diagnosis, prognosis evaluation, pathological classification, and staging [20–23]. In recent years, there has been growing interest in using radiomics to predict the expression levels of relevant disease molecular biomarkers and disease prognosis, showing potential promising clinical applications [24–27].

In this context, this study aims to investigate the correlation between CD38 expression levels and the survival status of patients with EOC, as well as explore preliminary potential mechanisms and immune microenvironment associations. Furthermore, we constructed two radiomics machine learning prediction models based on CT images to assess their feasibility for non-invasively predicting CD38 expression levels in the entire EOC tumour. These models also can indirectly reflect the prognosis and survival of patients with EOC.

2. Methods

2.1. Data acquisition

Transcriptome sequencing data, including clinical and follow-up data, were obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) to investigate the bioinformatics features associated with CD38 and its prognostic value for

patient survival. The inclusion criteria for the study were as follows: (1) patients with primary treatment ovarian serous cystadenocarcinoma (OSC), which is the main subtype of EOC [2]; (2) availability of RNA-seq data from the primary solid tumour. By contrast, the exclusion criteria were as follows: (1) cases that were not the primary treatment cases; (2) cases lacking survival status or follow-up time data; (3) data with a follow-up duration of less than 30 days; (4) cases lacking essential information, such as unknown pathological International Federation of Gynecology and Obstetrics (FIGO) staging, unknown pathological grading, or absence of RNA-seq results. Furthermore, cases with unclear identification of the solid primary tumour were also excluded. For survival analysis, the following seven variables were included as covariates: age (<60 years vs ≥ 60 years), chemotherapy (yes vs. no), FIGO stage (I/II vs. III vs. IV), lymphatic invasion (yes vs. no vs. unknown), tumour histological grade (G1/G2 vs. G3/G4), tumour residual disease (no macroscopic disease vs. 1–10 mm vs. ≥11 mm vs. unknown), and vascular invasion (yes vs. no vs. unknown).

The plain and enhanced abdominal CT imaging data, including clinical and follow-up data, were downloaded from The Cancer Imaging Archive (TCIA) database (https://www.cancerimagingarchive.net/). The inclusion criteria involved abdominal CT imaging data that intersected with the selected TCGA database cases. The exclusion criteria were applied to remove imaging data with inadequate image quality, defined as images with significant artefacts or poor resolution. The final data set including TCGA-TCIA intersecting data was used for extracting the radiomic features and constructing the prediction models.

2.2. CD38-related bioinformatics and survival analysis

The optimal cut-off value for the CD38 mRNA expression levels was calculated by identifying significant differences (corresponding to the minimum P-value) in the survival analysis using the "survminer" package in R (version 4.2.1). The data were then divided into high- and low-expression groups, which were labelled as $CD38^{high}$ and $CD38^{low}$, respectively. After processing through the Toil pipeline [28], the RNA data in TPM format for EOC from the TCGA dataset, along with the RNA data of corresponding normal tissue from the Genotype-tissue expression (GTEx) database, were downloaded and extracted from UCSC XENA (https://xenabrowser.net/datapages/). The "limma" package in R was used to compare differences in expression between the CD38 and EOC tissues and normal tissues.

The gene expression matrix of the EOC RNA-seq data was analysed using the ImmuCellAI database (http://bioinfo.life.hust.edu.cn/ImmuCellAI/#!/) to assess immune cell infiltration status. The differential immune cell infiltration between the $CD38^{high}$ and $CD38^{low}$ groups was compared. The "clusterProfiler" package in R was used to perform gene set enrichment analysis (GSEA) on the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Hallmark gene sets. A bar plot visualising the top 30 enriched pathways was generated. Additionally, the correlation between CD38 expression and apoptosis genes was also investigated.

The main outcome for clinical prognosis analysis was overall survival (OS). The "survival" package in R was used to plot evaluate Kaplan–Meier survival curves, followed by the evaluation of the survival rates of $CD38^{high}$ and $CD38^{low}$ groups. The difference in the survival rates between the two groups was assessed using the log-rank test. The "survival" package in R was also used to conduct univariate and multivariate Cox regression analyses, investigating the impact of various factors on the OS.

2.3. Images segmentation and radiomics feature extraction

The CT images were resampled to $1 \times 1 \times 1$ mm^3 pixels to achieve isotropic resolution. The images were then standardised to minimise discrepancies in grayscale values resulting from image acquisition on different machines.

Lesion segmentation was performed using 3D Slicer software (version 4.10.2; https://www.slicer.org/). A radiologist (H Liu) with 12 years of expertise in gynaecological imaging diagnosis, blinded to the CD38 expression levels and other clinical data, manually delineated the lesions layer by layer. Then, a random group of 10 cases was selected using the "random number table method" for secondary delineation by a senior radiologist (HJ Zhang) with 18 years of experience. The open-source "Pyradiomics" package (https://pyradiomics.readthedocs.io/en/latest) based on Python (version 3.7.4, https://www.python.org/downloads/) was employed to extract radiomic features. The extracted radiomic features were Z-score standardised using the "caret" package in R. The consistency of the radiomic features extracted from the volumes of interest (VOIs) delineated by the two radiologists was evaluated using the intraclass correlation coefficient (ICC) with the "irr" package in R. Features with an ICC ≥0.80 were selected for further selection.

2.4. Radiomics feature selection

Features with an ICC of 0.80 or higher were included. The maximum relevance minimum redundancy (mRMR) algorithm was used to select the 10 features with the highest relevance to the classification variable while minimising redundancy among features. Next, we utilised the stepwise regression algorithm with the Akaike information criterion (AIC) for further features selection. Finally, the selected features were included in the final set of radiomics features.

2.5. Construction and evaluation of prediction models

Two radiomics prediction models, logistic regression (LR) model and support vector machine (SVM) model, were constructed. The sigmoid function was used to map the linear combination of selected radiomics features to a value between 0 and 1, allowing for the establishment of a binary classification model for predicting CD38 expression. The "stats" package in R was used to establish the LR model.

On the other hand, the SVM algorithm utilises support vectors to find a hyperplane in high-dimensional space. This hyperplane

maximises the margin between the classification hyperplane and support vectors, thereby identifying the optimal decision boundary. The "caret" package in R was used to model the selected radiomics features using SVM.

The performance of the LR and SVM radiomics models was assessed using 5-fold cross-validation. The receiver operating characteristic (ROC) curve was plotted and area under the curve (AUC) was calculated to evaluate the diagnostic performance. A precision-recall (PR) curve was also generated to assess the discriminative and diagnostic ability, with the corresponding AUC values. Calibration curves were plotted and the Hosmer-Lemeshow goodness-of-fit test was used to evaluate the calibration degree of the predicted and observed results. Decision curve analysis (DCA) was performed to assess the clinical utility. The AUC values of the models before and after cross-validation were compared using the DeLong test. The analyses were conducted using the "pROC", "measures", "ResourceSelection", "rms", and "rmda" packages in R.

The LR and SVM radiomics models output the probability of CD38 expression level, which was then used as the radiomics score (Rad_score). Then, the difference in the Rad_scores between the $CD38^{high}$ and $CD38^{low}$ groups was evaluated.

2.6. Statistical analysis

Numerical data were reported as the mean ± standard deviation, while categorical data were presented as frequencies and percentages. The baseline characteristics of categorical variables were compared using the Chi-square test, whereas continuous variables were compared using the Wilcoxon test. The correlation between two factors was analysed using Spearman's correlation coefficient. The differences in quantitative data between two groups were compared using the Wilcoxon test. Statistical analysis was conducted using R software (version 4.2.1).

3. Results

3.1. Baseline characteristics

Based on the inclusion and exclusion criteria, data from a total of 333 patients were included from the TCGA database. The cut-off value for the CD38 expression level was calculated as 0.757, and according to this criterion, the patients were divided into $CD38^{high}$ group (145 cases) and $CD38^{low}$ group (188 cases). The baseline characteristics of the included patients are shown in Table 1. Except for tumour residual disease ($P = 0.024$), there were no significant differences in other clinical features between the two groups. After screening the TCIA database, a total of 56 TCGA-TCIA intersecting cases were ultimately included. An inclusion–exclusion flowchart is provided in Fig. 1.

3.2. Expression levels, immune infiltrates, and gene enrichment analysis of CD38

CD38 expression was significantly upregulated in EOC compared to normal tissues ($P < 0.01$), as shown in Fig. 2-A. Differences in

Table 1
Clinical characteristics of the EOC patients with high and low CD38 expression.

Variables	Total (n = 333)	Low (n = 188)	High (n = 145)	P
Age, n (%)				
<60	171(51)	100 (53)	71(49)	0.513
≥60	162(49)	88(47)	74(51)	
FIGO stage, n (%)				
I/II	18(5)	8(4)	10(7)	0.324
III	265(80)	148(79)	117(81)	
IV	50(15)	32(17)	18(12)	
Histologic grade, n (%)				
G1/G2	41(12)	28(15)	13(9)	0.143
G3/G4	292(88)	160(85)	132(91)	
Tumor residual disease, n (%)				
No macroscopic disease	58(18)	34(18)	24(17)	0.024
1–10 mm	161(48)	95(50)	66(46)	
≥11 mm	84(25)	50(27)	34(23)	
Unknown	30(9)	9(5)	21(14)	
Lymphatic invasion, n (%)				
Yes	90(27)	45(24)	45(31)	0.127
No	39(12)	27(14)	12(8)	
Unknown	204(61)	116(62)	88(61)	
Venous invasion, n (%)				
Yes	58(18)	30(16)	28(19)	0.510
No	31(9)	20(11)	11(8)	
Unknown	244(73)	138(73)	106(73)	
Chemotherapy, n (%)				
Yes	312(94)	177(94)	135(93)	0.871
No	21(6)	11(6)	10(7)	

the degree of immune cell infiltration were observed between the $CD38^{high}$ and $CD38^{low}$ groups. Specifically, the $CD38^{high}$ group showed a significant increase in the infiltration of some types of immune cells, such as Natural Killer (NK) cells, dendritic cells (DC), and gamma-delta T (γδT) cells compared to the $CD38^{low}$ group ($P < 0.001$), as shown in Fig. 2-B.

The results of GSEA revealed significant enrichment of differentially expressed genes between the $CD38^{high}$ and $CD38^{low}$ groups in multiple signalling pathways, mainly including the mitogen-activated protein kinase (MAPK) signalling pathway, chemokine signalling pathway, apoptosis, and PI3K-AKT-mTOR signalling, as shown in Fig. 2-C and D. CD38 expression was significantly positively correlated ($P < 0.001$) with CSF2RB, PIK3CG, and FASLG genes, as shown in Fig. 3.

3.3. CD38-related survival analysis

The Kaplan–Meier survival curves for the $CD38^{high}$ and $CD38^{low}$ groups are provided in Fig. 4-A. The median survival time was 52.77 months for the $CD38^{high}$ group and 42.13 months for $CD38^{low}$ group, indicating a significant association between high CD38 expression and prolonged OS ($P < 0.001$).

The results of the univariate Cox regression analysis indicated that high levels of CD38 expression (HR = 0.554, 95 % CI: 0.415–0.740, $P < 0.001$) and chemotherapy (HR = 0.470, 95 % CI: 0.288–0.767, $P = 0.003$) were protective factors for OS. Similarly, the results of the multivariate Cox regression analysis demonstrated that high levels of CD38 expression (HR = 0.540, 95 % CI: 0.400–0.730, $P < 0.001$) and chemotherapy (HR = 0.377, 95 % CI: 0.226–0.627, $P < 0.001$) were independent protective factors for OS, as shown in Fig. 4-B. In addition, we performed COX subgroup analysis on the main variable CD38, which confirmed that there was no significant interaction between the covariates and the association of CD38 and OS. Further details are provided in Fig. 5.

3.4. Radiomics feature extraction and selection

The 56 cases of TCGA-TCIA were divided into the $CD38^{high}$ (n = 29) and $CD38^{low}$ (n = 27) groups based on the cut-off value of 0.757. A total of 107 radiomics features were extracted from manually delineated tumour regions (an example of VOI delineation is provided in Fig. 6) and standardised. The median value of the ICC was 0.963, with 101 features having an ICC ≥0.80, five features with 0.5 ≤ ICC ≤0.8, and one feature with ICC <0.5. Features with an ICC ≥0.80 were further selected for subsequent feature selection.

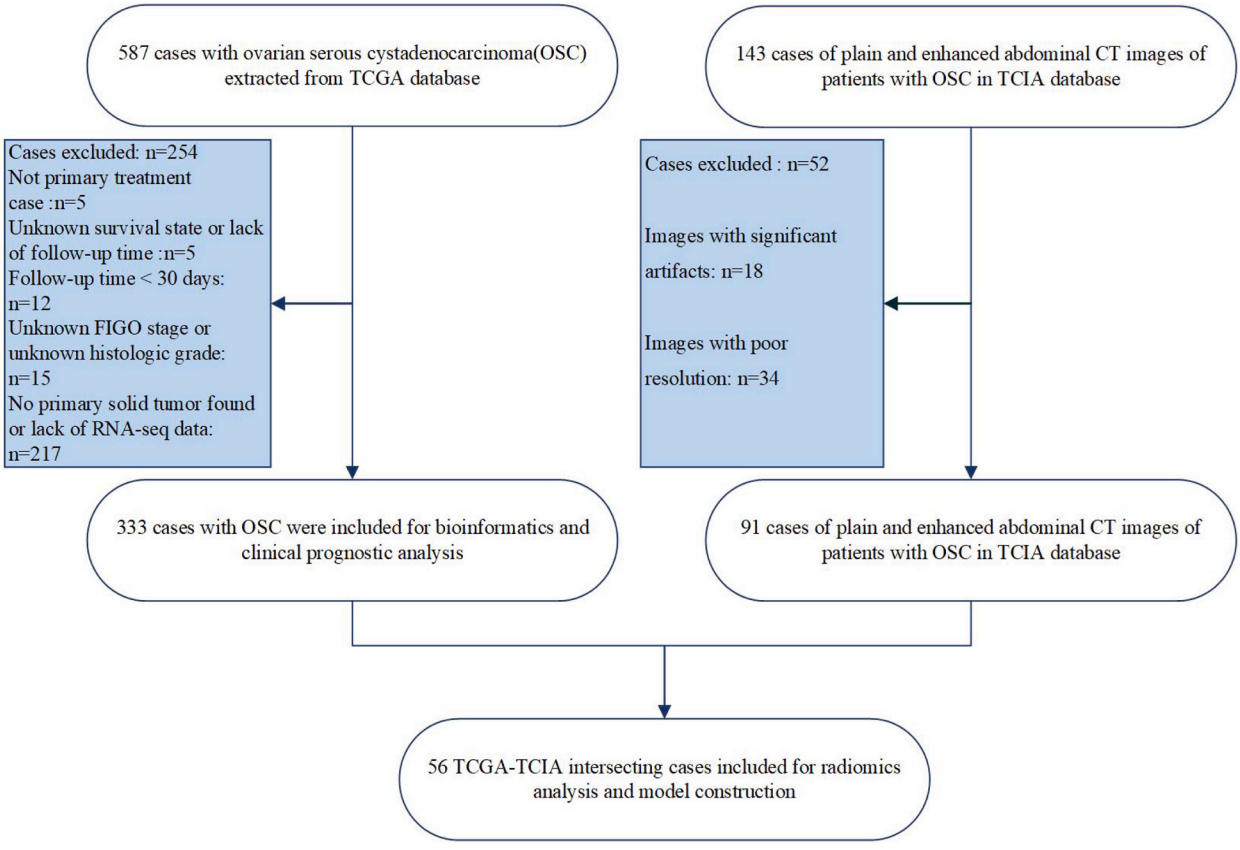


Fig. 1. Inclusion–exclusion flowchart of patients with epithelial ovarian cancer (EOC) from The Cancer Genome Atlas (TCGA) database and The Cancer Imaging Archive (TCIA) database.

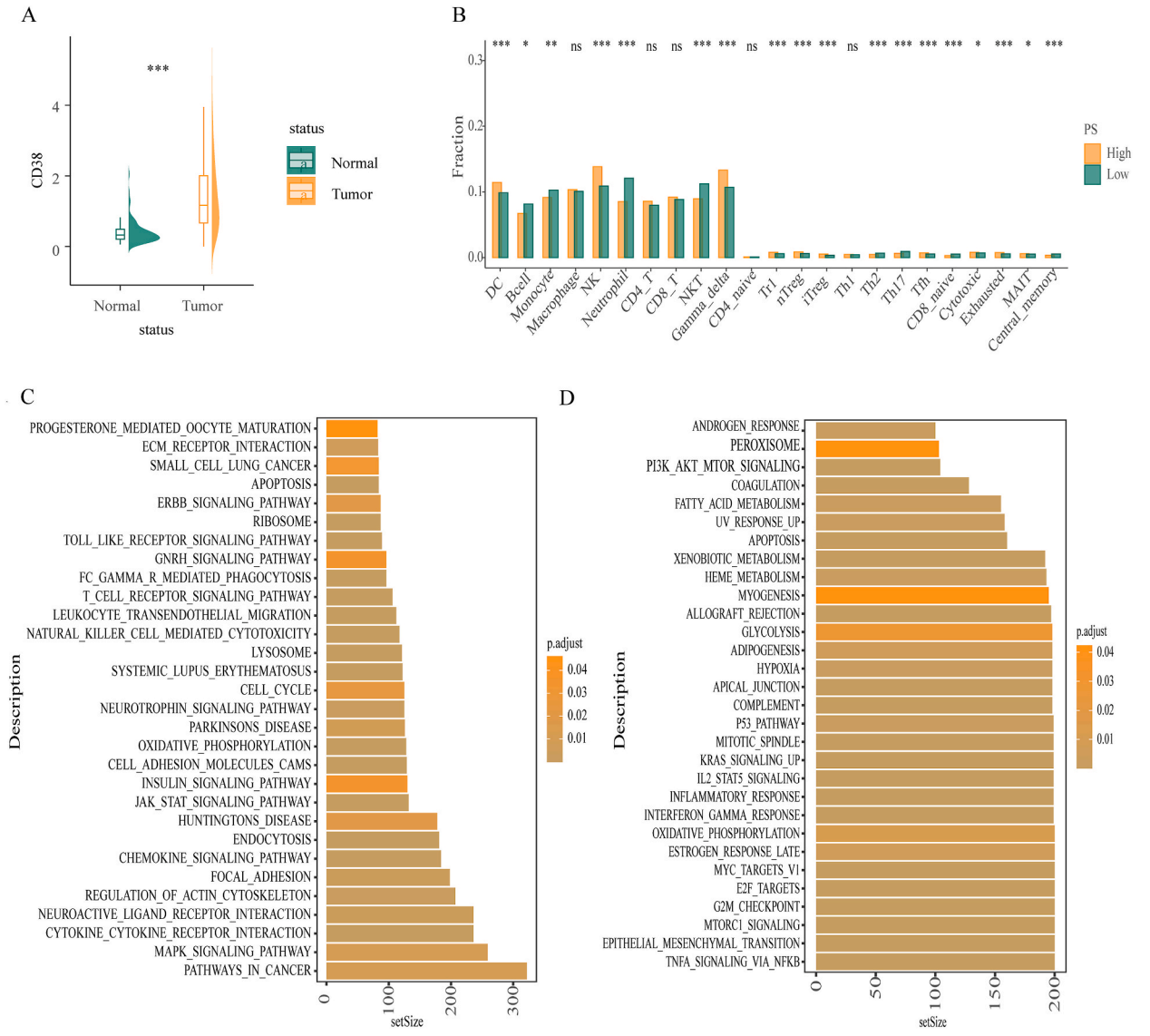


Fig. 2. CD38 expression, immune cell infiltration, and gene set enrichment analysis (GSEA) of epithelial ovarian cancer (EOC). **(A)** CD38 expression in tumour tissue and normal tissue based on TCGA-OV database and GTEx database. **(B)** Differences in immune cell infiltration between $CD38^{high}$ and $CD38^{low}$ groups based on the TCGA-OV database. **(C)** GSEA of KEGG gene set associated with CD38 expression. **(D)** GSEA on Hallmark gene set associated with CD38 expression; ns, not significant; *P < 0.05, **P < 0.01, ***P < 0.001.

Finally, five radiomics features, including glcm_ldn, glrlm_RunLenthNonUniformity, gldm_DependenceVariance, glszm_LargeAreaHighGrayLevelEmphasis, and gldm_SmallDependenceLowGrayLevelEmphasis, were selected by mRMR and stepwise regression.

3.5. Construction and evaluation of LR and SVM models

In the LR model, the overall importance of the selected radiomic features mentioned above were ranked as 1.421, 1.346, 1.675, 0.864, and 2.346, respectively. The formula for predicting the probability of the CD38 expression level, represented by the Rad_score, was as follows:

Rad_score = 0.220–0.547 × (glcm_ldn) − 0.706 × (glrlm_RunLenthNonUniformity) − 0.566 × (gldm_DependenceVariance) + 1.787 × (glszm_LargeAreaHighGrayLevelEmphasis) − 1.181 × (gldm_SmallDependenceLowGrayLevelEmphasis)

The SVM model was also constructed using the aforementioned five radiomics features. The importance of each radiomics feature in the model was 0.557, 0.512, 0.567, 0.502, and 0.572, respectively.

Next, the predictive performance of the LR and SVM models was evaluated. The LR model achieved an AUC value of 0.739 (95 % CI:

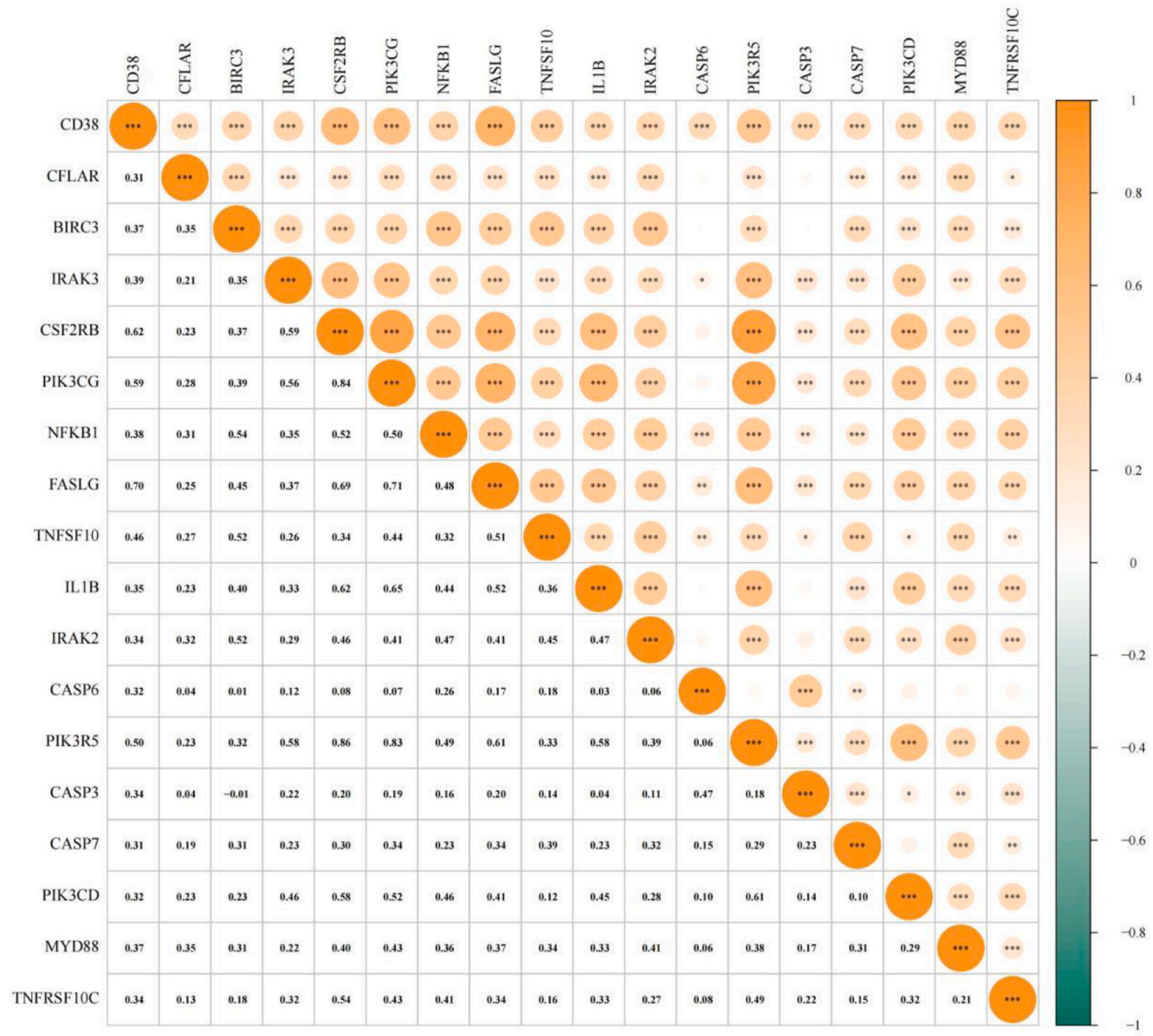


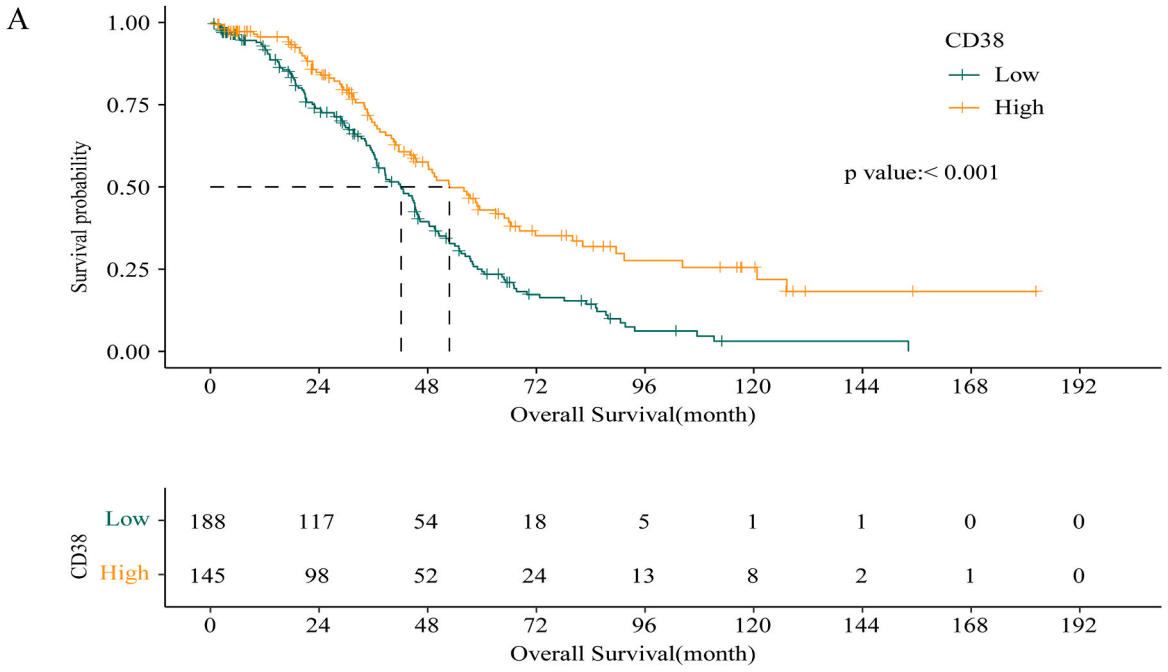
Fig. 3. Heat map of correlation analysis between CD38 expression and apoptosis genes with coefficients greater than 0.3. *P < 0.05, **P < 0.01, ***P < 0.001.

0.609–0.870) for the ROC curve, and after 5-fold cross-validation, the ROC-AUC of the LR model was 0.732 (95 % CI:0.595–0.868), as shown in Fig. 7-A and B. The SVM model had an AUC value of 0.741 (95 % CI: 0.608–0.874) for the ROC curve, and a 5-fold cross-validated ROC-AUC of 0.700 (95 % CI: 0.555–0.835), as shown in Fig. 8-A and B. Hosmer-Lemeshow goodness-of-fit test demonstrated comparatively good consistency between the predicted probabilities and the actual values for both the LR (P = 0.838) and SVM (P = 0.074) models. The corresponding calibration curves are shown in Fig. 7-C and 8-C. The AUC for the PR curve was 0.760 for the LR model (Fig. 7-D) and 0.721 for the SVM model (Fig. 8-D). The DCA curves indicated that if the risk threshold probability was between 30 % and 80 %, the two model both achieved a superior net benefit (Fig. 7-E and 8-E). The DeLong test indicated no statistically significant difference between the AUC values before cross-validation and after cross-validation in the two models (P = 0.961 in LR model and P = 0.703 in SVM model), indicating the stability of the model fit.

The LR and SVM models both output the probability of predicting the CD38 expression level, referred to as Rad_score. A significant difference was observed in the distribution of Rad_score between the $CD38^{high}$ and $CD38^{low}$ groups (P < 0.01), wherein the $CD38^{high}$ group exhibited a higher Rad_score (Fig. 7-F and 8-F).

4. Discussion

The results demonstrated that: (1) CD38 is highly expressed in EOC tissues; the expression level of CD38 was associated with



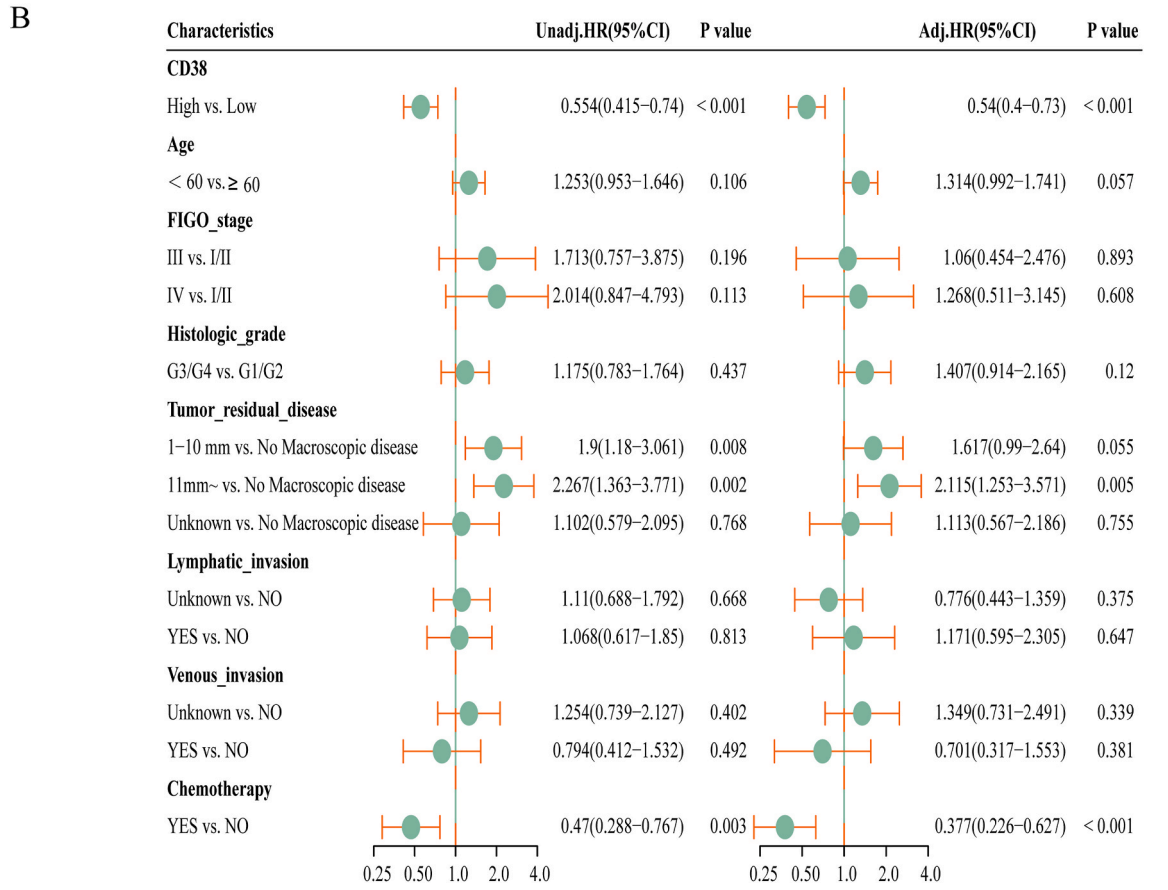


Fig. 4. Survival analysis of TCGA-OV cohort. **(A)** The Kaplan–Meier curve of overall survival in the CD38^{high} and CD38^{low} groups. **(B)** Univariate and multivariate Cox regression analyses.

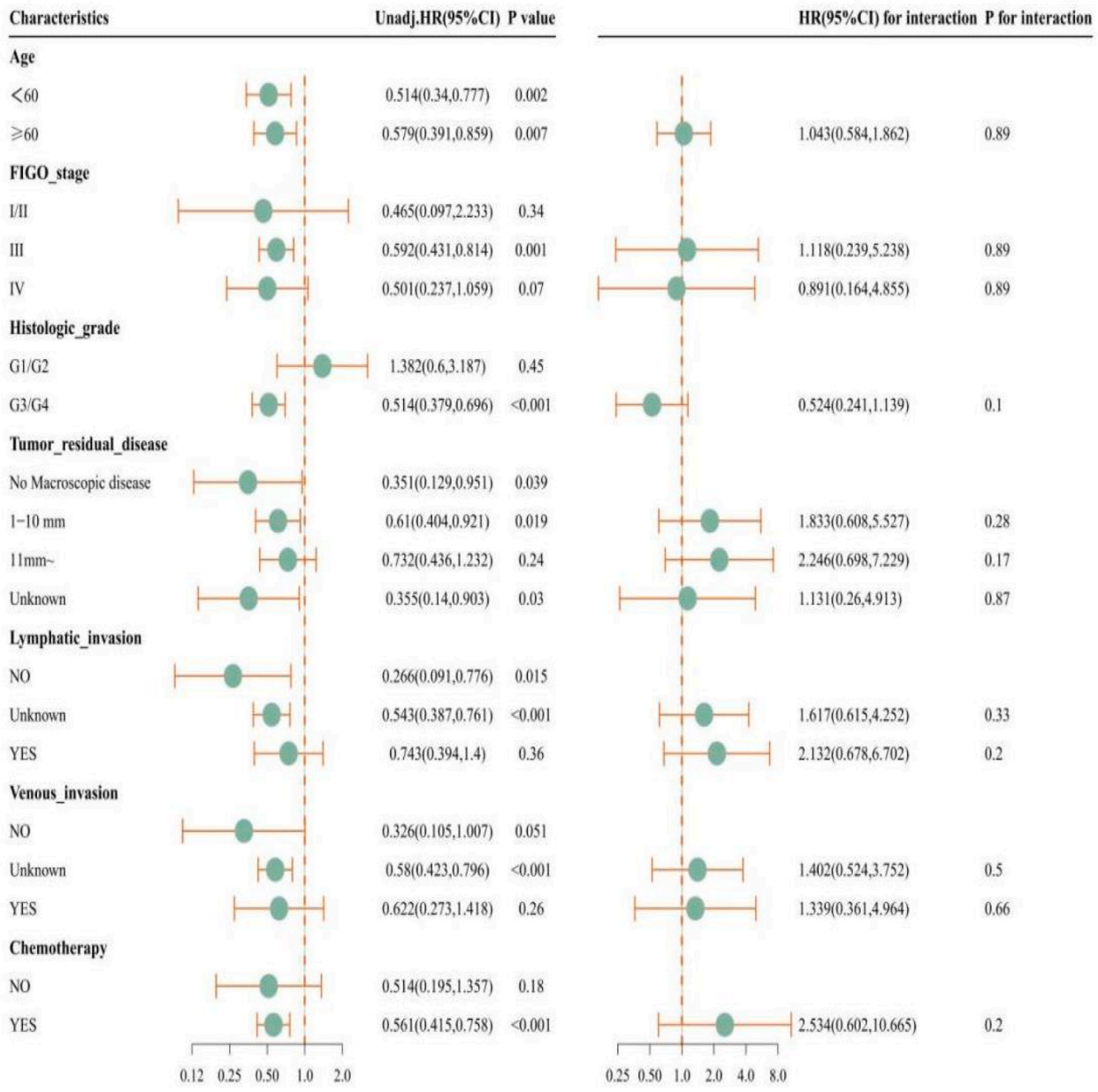


Fig. 5. COX subgroup analysis on the main variable CD38.

immune cell infiltration, some tumour regulatory pathways, and microenvironment regulation; (2) high CD38 expression is an independent protective factor of OS in EOC patients; (3) the two machine learning models based on CT radiomics features were able to predict the expression levels of CD38.

EOC exhibits a high degree of heterogeneity, which contributes to marked differences in the patients' prognoses [29]. In recent years, research on molecular mechanisms related to tumour metabolism regulation and the tumour microenvironment has become a hot topic, with the objective of accurately predicting the prognosis of EOC. Some relevant molecules may have potential value as prognostic biomarkers [30–33].

Our findings suggest that CD38 could serve as a biomarker for prognosis and survival. Alongside previous research [14,34], this study demonstrated the protective effect of high levels of CD38 expression on improvements in the survival of EOC patients. A possible correlation with immune cell infiltration and signalling pathways has also been investigated, with the $CD38^{high}$ group exhibiting an increased infiltration of NK cells [35] and γδT cells [36] with tumour-killing ability and immune regulation, as well as DC [37] that induce a $CD8^{+}$ T cell-mediated immune response [34,38]. The differences observed between the $CD38^{high}$ and $CD38^{low}$ groups primarily involved signalling pathways regulating cell proliferation, growth, differentiation, apoptosis, and tumour metabolism. CD38

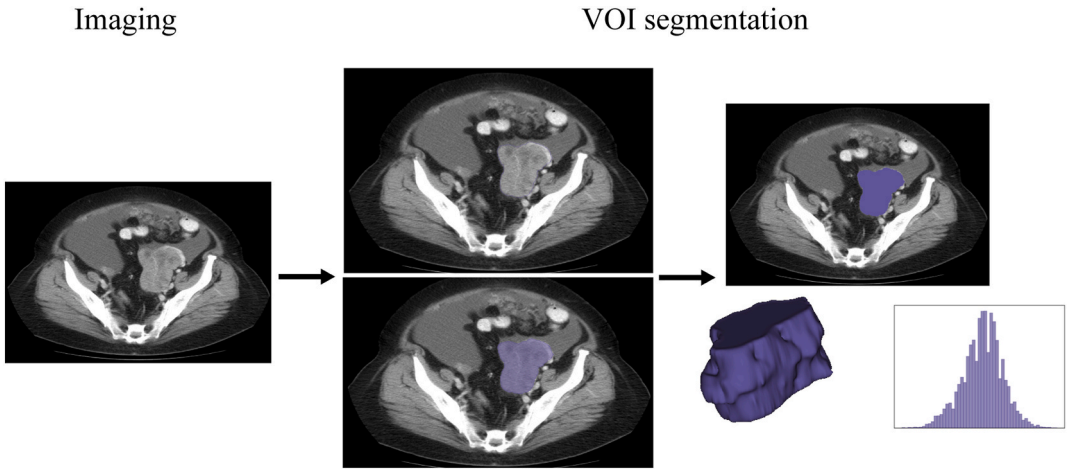


Fig. 6. An example of VOI delineation processing.

has been confirmed to regulate and activate several signalling pathways, such as PI3_AKT_MTOR_SIGNALING [12].

In recent years, machine learning methods based on radiomics have been applied in various areas, including tumour diagnosis, differential diagnosis, staging, treatment response, and survival prediction, showcasing their promising diagnostic and predictive capabilities [39–43]. Several studies have endeavoured to develop radiomics models utilising CT, MRI, and other imaging modalities to predict biomarker expression levels in different tumour tissues. For instance, these models have successfully predicted the expression levels of phosphorylated β-arrestin in hepatocellular carcinoma [44] and low-density lipoprotein receptor-related protein-1 in locally advanced rectal cancer [45], as well as C–C motif chemokine receptor 5 [25] and programmed cell death protein 1 [24] in OC. Therefore, radiomics has shown sufficient value in the non-invasive prediction of biomarker expression.

In the present study, two radiomics models for CD38 expression prediction based on CT images of EOC patients were established. The AUCs of ROC both revealed the fine predictive ability of the LR and SVM models. The PR curve also highlight the adequate discriminative ability of these two models. However, the LR model showed a better predictive performance than the SVM model. Moreover, both models exhibited similar AUC values in both the training and validation sets, indicating consistent and stable performances. Additionally, the resulting calibration curves demonstrated a satisfactory consistency between the observed results and their predicted values, with the DCA curves indicating a potential clinical net benefit for their application.

To the best of our knowledge, this study may represent the first attempt to develop predictive radiomics models for assessing CD38 expression in patients with EOC based on CT imaging. CT imaging is widely utilised in clinical practice and the images are easily obtainable. The radiomics approach enables the extraction of quantitative features from conventional images, providing abundant objective information for further analysis. Therefore, the CT-based radiomics models offer a simple, non-invasive, and cost-effective method to predict the expression level of CD38. These models have the potential to indirectly reflect the prognosis and survival outcomes of EOC patients. Furthermore, based on previous research, CD38, as a well-known therapeutic target for multiple myeloma [10], has been suggested as a potential target for solid tumours [17,18]. In the future, by utilising the radiomics model, EOC patients with high levels of CD38 expression may potentially receive individualised targeted therapies at an early stage, leading to improved prognosis and survival. Furthermore, future investigations could involve the non-invasive and dynamic monitoring of patients through the analysis of sequential imaging data before and after treatment.

This study has several limitations. Firstly, all the original images were obtained from the TCIA public database, which may introduce inherent variations between the images and potentially impact subsequent image analysis. Secondly, the relatively small sample size used for model construction in this retrospective study brings about a risk of overfitting. Furthermore, the effectiveness of the model was evaluated using internal validation alone, lacking an independent external validation set. Thirdly, the cases included in our analysis were all of the OSC subtype; therefore, the utility of the constructed model in other EOC subtypes remains to be elucidated. Moreover, the manual delineation of VOIs is subjective, introducing potential biases that can affect subsequent analysis. Therefore, it is necessary to conduct multicentre studies with larger sample sizes. Additionally, incorporating more objective radiomics practice methods, such as automated or semi-automated VOI delineation, could enhance the accuracy and reliability of the analysis.

5. Conclusion

High levels of CD38 expression can improve OS in EOC patients. In this study, we developed CT-based radiomics machine learning models using LR and SVM algorithms to predict the expression levels of CD38 in EOC tissues. The resulting models, especially the LR model, exhibited a fine predictive performance, providing potential for individualised non-invasive prognosis assessment and clinical decision-making in patients with EOC.

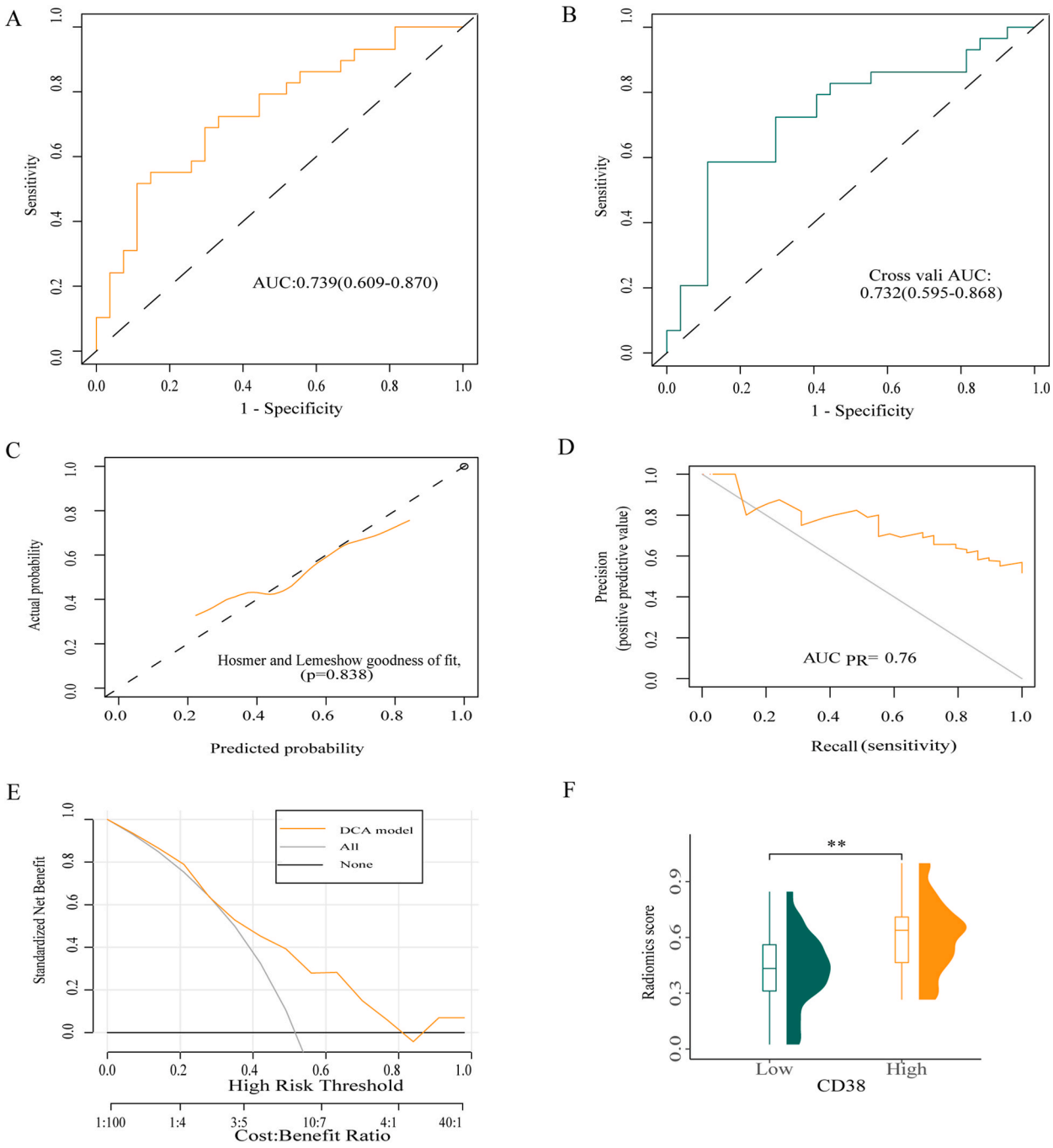


Fig. 7. Performance of the logistic regression (LR) radiomics model for predicting CD38 expression level. **(A)** Receiver operating characteristic (ROC) curve of the LR model in the training set. **(B)** ROC curve of the LR model in the internal validation set with 5-fold cross-validation. **(C)** Calibration curve of the LR model. **(D)** Precision-recall (PR) curve of the LR model. **(E)** Decision curve analysis (DCA) for the LR model. **(F)** The distribution of the radiomics score (Rad_score) of the LR model between the $CD38^{high}$ and $CD38^{low}$ groups. **P < 0.01.

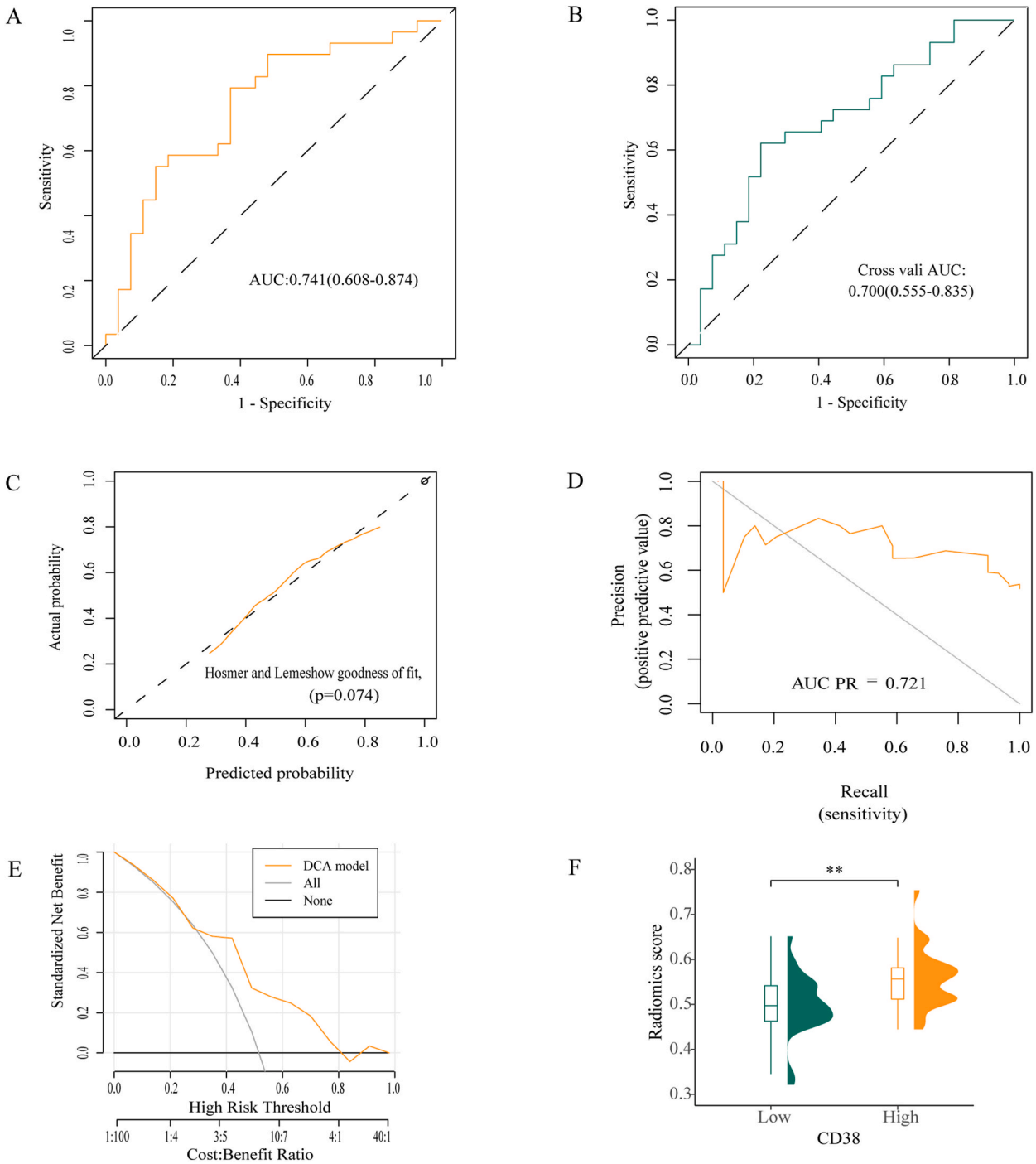


Fig. 8. Performance of the support vector machine (SVM) radiomics model for predicting the level of CD38 expression. **(A)** ROC curve of the SVM model in the training set. **(B)** ROC curve of the SVM model in the internal validation set with 5-fold cross-validation. **(C)** Calibration curves of the SVM model. **(D)** PR curve of the SVM model. **(E)** DCA for the SVM model. **(F)** The distribution of the Rad_score of the SVM model between the $CD38^{high}$ and $CD38^{low}$ groups. **P < 0.01.

Funding

This study was funded by Jiaxing Science and Technology Plan Project (2021AD30028).

Ethics approval

Review and/or approval by an ethics committee was not needed for this study as the original research within the TCIA and TCGA databases had already undergone ethical approval and obtained informed consent from the participants. Furthermore, the databases have anonymised sensitive patient information, enabling the data in these public databases to be openly accessible.

Consent for publication

Not required.

Data availability statement

All data generated or analysed during this study are included in this published article and its supplementary information files. The datasets for this study were downloaded from TCIA (https://www.cancerimagingarchive.net/) and TCGA (https://portal.gdc.cancer.gov/) database.

CRediT authorship contribution statement

Yuan Yao: Writing – original draft, Methodology, Funding acquisition, Formal analysis. **Haijin Zhang:** Software, Methodology, Funding acquisition. **Hui Liu:** Investigation, Data curation. **Chendi Teng:** Investigation, Formal analysis. **Xuan Che:** Writing – review & editing, Project administration, Conceptualization. **Wei Bian:** Methodology. **Wenting Zhang:** Data curation. **Zhifeng Wang:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the study participants for permitting us to use their personal data. We are also sincerely thankful for the support and assistance provided by the pathologists W.G. Zhou and J.Q. Xu.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32910.

References

[1] S. Lheureux, M. Braunstein, A.M. Oza, Epithelial ovarian cancer: evolution of management in the era of precision medicine, CA Cancer J Clin 69 (4) (2019) 280–304, https://doi.org/10.3322/caac.21559.
[2] S. Lheureux, et al., Epithelial ovarian cancer, Lancet 393 (10177) (2019) 1240–1253, https://doi.org/10.1016/S0140-6736(18)32552-2.
[3] L. Kuroki, S.R. Guntupalli, Treatment of epithelial ovarian cancer, BMJ 371 (2020) m3773, https://doi.org/10.1136/bmj.m3773.
[4] R.L. Siegel, et al., Cancer statistics, 2023, CA Cancer J Clin 73 (1) (2023) 17–48, https://doi.org/10.3322/caac.21763.
[5] K. Gaitskell, et al., Ovarian cancer survival by stage, histotype, and pre-diagnostic lifestyle factors, in the prospective UK Million Women Study, Cancer Epidemiol 76 (2022) 102074, https://doi.org/10.1016/j.canep.2021.102074.
[6] Y. Rong, L. Li, Early clearance of serum HE4 and CA125 in predicting platinum sensitivity and prognosis in epithelial ovarian cancer, J. Ovarian Res. 14 (1) (2021) 2, https://doi.org/10.1186/s13048-020-00759-9.
[7] S. Rizzo, et al., Imaging before cytoreductive surgery in advanced ovarian cancer patients, Int. J. Gynecol. Cancer 30 (1) (2020) 133–138, https://doi.org/10.1136/ijgc-2019-000819.
[8] W. Li, et al., CD38: an important regulator of T cell function, Biomed. Pharmacother. 153 (2022) 113395, https://doi.org/10.1016/j.biopha.2022.113395.
[9] L. Gao, et al., Evolving roles of CD38 metabolism in solid tumour microenvironment, Br. J. Cancer 128 (4) (2023) 492–504, https://doi.org/10.1038/s41416-022-02052-6.
[10] A. Gozzetti, et al., Anti CD38 monoclonal antibodies for multiple myeloma treatment, Hum Vaccin Immunother 18 (5) (2022) 2052658, https://doi.org/10.1080/21645515.2022.2052658.
[11] M.J. Zhang, et al., CD38 triggers inflammasome-mediated pyroptotic cell death in head and neck squamous cell carcinoma, Am. J. Cancer Res. 10 (9) (2020) 2895–2908.
[12] S. Liao, et al., CD38 is involved in cell energy metabolism via activating the PI3K/AKT/mTOR signaling pathway in cervical cancer cells, Int. J. Oncol. 57 (1) (2020) 338–354, https://doi.org/10.3892/ijo.2020.5040.

[13] C. Guo, et al., CD38 in advanced prostate cancers, Eur. Urol. 79 (6) (2021) 736–746, https://doi.org/10.1016/j.eururo.2021.01.017.
[14] Y. Zhu, et al., CD38 predicts favorable prognosis by enhancing immune infiltration and antitumor immunity in the epithelial ovarian cancer microenvironment, Front. Genet. 11 (2020) 369, https://doi.org/10.3389/fgene.2020.00369.
[15] K.U. Choi, et al., Differences in immune-related gene expressions and tumor-infiltrating lymphocytes according to chemotherapeutic response in ovarian high-grade serous carcinoma, J. Ovarian Res. 13 (1) (2020) 65, https://doi.org/10.1186/s13048-020-00667-y.
[16] J. Zhou, et al., Clinical significance of CD38 and CD101 expression in PD-1(+)CD8(+) T cells in patients with epithelial ovarian cancer, Oncol. Lett. 20 (1) (2020) 724–732, https://doi.org/10.3892/ol.2020.11580.
[17] P.A. Zucali, et al., Targeting CD38 and PD-1 with isatuximab plus cemiplimab in patients with advanced solid malignancies: results from a phase I/II open-label, multicenter study, J Immunother Cancer 10 (1) (2022), https://doi.org/10.1136/jitc-2021-003697.
[18] Y. Jiao, et al., CD38: targeted therapy in multiple myeloma and therapeutic potential for solid cancers, Expert Opin Investig Drugs 29 (11) (2020) 1295–1308, https://doi.org/10.1080/13543784.2020.1814253.
[19] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2) (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.
[20] K. Bera, et al., Predicting cancer outcomes with radiomics and artificial intelligence in radiology, Nat. Rev. Clin. Oncol. 19 (2) (2022) 132–146, https://doi.org/10.1038/s41571-021-00560-7.
[21] Z. Feng, et al., CT radiomics to predict macrotrabecular-massive subtype and immune status in hepatocellular carcinoma, Radiology 307 (1) (2023) e221291, https://doi.org/10.1148/radiol.221291.
[22] G. Li, et al., An MRI radiomics approach to predict survival and tumour-infiltrating macrophages in gliomas, Brain 145 (3) (2022) 1151–1161, https://doi.org/10.1093/brain/awab340.
[23] Y. Qi, T. Zhao, M. Han, The application of radiomics in predicting gene mutations in cancer, Eur. Radiol. 32 (6) (2022) 4014–4024, https://doi.org/10.1007/s00330-021-08520-6.
[24] L. Gao, et al., Radiomic model to predict the expression of PD-1 and overall survival of patients with ovarian cancer, Int Immunopharmacol 113 (Pt A) (2022) 109335, https://doi.org/10.1016/j.intimp.2022.109335.
[25] S. Wan, et al., CT-based machine learning radiomics predicts CCR5 expression level and survival in ovarian cancer, J. Ovarian Res. 16 (1) (2023) 1, https://doi.org/10.1186/s13048-022-01089-8.
[26] A.K. Meissner, et al., Radiomics for the non-invasive prediction of PD-L1 expression in patients with brain metastases secondary to non-small cell lung cancer, J. Neuro Oncol. 163 (3) (2023) 597–605, https://doi.org/10.1007/s11060-023-04367-7.
[27] T. Ramtohul, et al., Multiparametric MRI and radiomics for the prediction of HER2-zero, -low, and -positive breast cancers, Radiology 308 (2) (2023) e222646, https://doi.org/10.1148/radiol.222646.
[28] J. Vivian, et al., Toil enables reproducible, open source, big biomedical data analyses, Nat. Biotechnol. 35 (4) (2017) 314–316, https://doi.org/10.1038/nbt.3772.
[29] F. Reid, et al., The World Ovarian Cancer Coalition Every Woman Study: identifying challenges and opportunities to improve survival and quality of life, Int. J. Gynecol. Cancer 31 (2) (2021) 238–244, https://doi.org/10.1136/ijgc-2019-000983.
[30] S. Banerjee, et al., Targeting NaPi2b in ovarian cancer, Cancer Treat Rev. 112 (2023) 102489, https://doi.org/10.1016/j.ctrv.2022.102489.
[31] A.W. Gahlawat, et al., A novel circulating miRNA panel for non-invasive ovarian cancer diagnosis and prognosis, Br. J. Cancer 127 (8) (2022) 1550–1556, https://doi.org/10.1038/s41416-022-01925-0.
[32] L. Ge, et al., Plasma circRNA microarray profiling identifies novel circRNA biomarkers for the diagnosis of ovarian cancer, J. Ovarian Res. 15 (1) (2022) 58, https://doi.org/10.1186/s13048-022-00988-0.
[33] X. Xi, et al., CDC20 is a novel biomarker for improved clinical predictions in epithelial ovarian cancer, Am. J. Cancer Res. 12 (7) (2022) 3303–3317.
[34] L. Li, et al., Identification of CD8(+) T cell related biomarkers in ovarian cancer, Front. Genet. 13 (2022) 860161, https://doi.org/10.3389/fgene.2022.860161.
[35] M. Pugh-Toole, et al., Natural killer cells: the missing link in effective treatment for high-grade serous ovarian carcinoma, Curr. Treat. Options Oncol. 23 (2) (2022) 210–226, https://doi.org/10.1007/s11864-021-00929-x.
[36] J. Saura-Esteller, et al., Gamma delta T-cell based cancer immunotherapy: past-present-future, Front. Immunol. 13 (2022) 915837, https://doi.org/10.3389/fimmu.2022.915837.
[37] L. Rob, et al., Safety and efficacy of dendritic cell-based immunotherapy DCVAC/OvCa added to first-line chemotherapy (carboplatin plus paclitaxel) for epithelial ovarian cancer: a phase 2, open-label, multicenter, randomized trial, J Immunother Cancer 10 (1) (2022), https://doi.org/10.1136/jitc-2021-003190.
[38] H. Chang, et al., Construction of a macrophage infiltration regulatory network and related prognostic model of high-grade serous ovarian cancer, J Oncol 2021 (2021) 1331031, https://doi.org/10.1155/2021/1331031.
[39] M. Avanzo, et al., Radiomics and deep learning in lung cancer, Strahlenther. Onkol. 196 (10) (2020) 879–887, https://doi.org/10.1007/s00066-020-01625-9.
[40] Y. Yu, et al., Magnetic resonance imaging radiomics predicts preoperative axillary lymph node metastasis to support surgical decisions and is associated with tumor microenvironment in invasive breast cancer: a machine learning, multicenter study, EBioMedicine 69 (2021) 103460, https://doi.org/10.1016/j.ebiom.2021.103460.
[41] R. Autorino, et al., Radiomics-based prediction of two-year clinical outcome in locally advanced cervical cancer patients undergoing neoadjuvant chemoradiotherapy, Radiol. Med. 127 (5) (2022) 498–506, https://doi.org/10.1007/s11547-022-01482-9.
[42] K.E. Fasmer, et al., Whole-volume tumor MRI radiomics for prognostic modeling in endometrial cancer, J Magn Reson Imaging 53 (3) (2021) 928–937, https://doi.org/10.1002/jmri.27444.
[43] A. Bhandari, et al., CT-based radiomics for differentiating renal tumours: a systematic review, Abdom Radiol (NY) 46 (5) (2021) 2052–2063, https://doi.org/10.1007/s00261-020-02832-9.
[44] F. Che, et al., Radiomics signature: a potential biomarker for beta-arrestin1 phosphorylation prediction in hepatocellular carcinoma, World J. Gastroenterol. 28 (14) (2022) 1479–1493, https://doi.org/10.3748/wjg.v28.i14.1479.
[45] Z. Li, et al., DCE-MRI radiomics models predicting the expression of radioresistant-related factors of LRP-1 and survivin in locally advanced rectal cancer, Front. Oncol. 12 (2022) 881341, https://doi.org/10.3389/fonc.2022.881341.