cancers

MDPI

Article

Prognostic Value and Quantitative CT Analysis in RANKL Expression of Spinal GCTB in the Denosumab Era: A Machine Learning Approach

Qizheng Wang [1], Yongye Chen [1], Siyuan Qin [1], Xiaoming Liu [2,3], Ke Liu [1], Peijin Xin [1], Weili Zhao [1], Huishu Yuan [1] and Ning Lang [1,*]

[1] Department of Radiology, Peking University Third Hospital, Beijing 100191, China
[2] Department of Research and Development, United Imaging Intelligence (Beijing) Co., Ltd., Yongteng North Road, Haidian District, Beijing 100089, China
[3] Beijing United Imaging Research Institute of Intelligent Imaging, Yongteng North Road, Haidian District, Beijing 100089, China
* Correspondence: langning800129@126.com; Tel.: +86-01082265109; Fax: +86-860106900690

Simple Summary: Prognostic assessment of giant cell tumor of bone (GCTB) is an ongoing challenge in the treatment and management of bone tumors. Recurrence rates of spinal GCTB are higher compared to GCTB in other bone sites, presumably due to a more aggressive pathology and/or the conservative surgery performed to spare the spinal cord nerve function and decrease postoperative complications. A more accurate prognosis of GCTB will help to inform the choice of treatment methods. This retrospective study investigated prognosis-related molecular markers in spinal GCTB, including RANKL (target of denosumab), focusing on using machine learning analysis based on pre-operative CT to evaluate RANKL status, which may facilitate the selection of better disease management strategies.

Abstract: The receptor activator of the nuclear factor kappa B ligand (RANKL) is the therapeutic target of denosumab. In this study, we evaluated whether radiomics signature and machine learning analysis can predict RANKL status in spinal giant cell tumors of bone (GCTB). This retrospective study consisted of 107 patients, including a training set (n = 82) and a validation set (n = 25). Kaplan-Meier survival analysis was used to validate the prognostic value of RANKL status. Radiomic feature extraction of three heterogeneous regions (VOI^{entire}, VOI^{edge}, and VOI^{core}) from pretreatment CT were performed. Followed by feature selection using Selected K Best and least absolute shrinkage and selection operator (LASSO) analysis, three classifiers (random forest (RF), support vector machine, and logistic regression) were used to build models. The area under the curve (AUC), accuracy, F1 score, recall, precision, sensitivity, and specificity were used to evaluate the models' performance. Classification of 75 patients with eligible follow-up based on RANKL status resulted in a significant difference in progression-free survival (p = 0.035). VOI^{core}-based RF classifier performs best. Using this model, the AUCs for the training and validation cohorts were 0.880 and 0.766, respectively. In conclusion, a machine learning approach based on CT radiomic features could discriminate prognostically significant RANKL status in spinal GCTB, which may ultimately aid clinical decision-making.

Keywords: giant cell tumor of bone; prognosis; tomography; RANKL; radiomics

check for updates

Citation: Wang, Q.; Chen, Y.; Qin, S.; Liu, X.; Liu, K.; Xin, P.; Zhao, W.; Yuan, H.; Lang, N. Prognostic Value and Quantitative CT Analysis in RANKL Expression of Spinal GCTB in the Denosumab Era: A Machine Learning Approach. *Cancers* **2022**, *14*, 5201. https://doi.org/10.3390/cancers14215201

Academic Editors: Ramses Forsyth and Pancras Hogendoorn

Received: 17 August 2022
Accepted: 13 October 2022
Published: 23 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Over the last few years, our knowledge of the receptor activator of nuclear factor kappa-B (RANK)/RANK ligand (RANKL) pathway has expanded, and targeting RANKL is being evaluated as anticancer therapy [1,2]. Overexpression of RANKL increases the risk of metastasis in giant cell tumors of bone [3]. However, the effects of RANKL levels on the long-term survival of patients with spinal GCTB have not yet been elucidated.

Denosumab, a monoclonal antibody against RANKL, is a promising treatment option for patients with spinal giant cell tumor of bone (GCTB), and has been approved by the United States Food and Drug Administration (FDA) [4]. Spinal GCTB is usually of the high preoperative stage, and denosumab is used as a targeted therapy for tumors that are difficult to operate upon [5,6]. However, denosumab treatment may lead to extensive recurrence upon treatment withdrawal, and denosumab has been suggested to be used with caution [7–9]. Identifying patients who can benefit from denosumab treatment can be helpful in treatment decisions.

The rapid advances in immuno-oncology have revolutionized cancer treatment [10]. As cancer patients respond differently to immunotherapy, the use of targeted drugs is not a blanket statement. An accurate assessment of preoperative RANKL levels may also help clinicians and patients make treatment decisions, including the determination of whether drug therapy, extended resection, postoperative radiotherapy, intensive follow-up, and reasonable clinical expectations. Although tissue biopsy before surgery can be used for disease detection, aside from invasive risk, a small amount of biopsy tissue cannot be used enough to evaluate the entire tumor heterogeneity, and making assessments based solely on small tissue biopsies can lead to bias in diagnosis and evaluation [11].

Imaging is pivotal in assessing tumor heterogeneity. While X-ray is the primary imaging modality for the diagnosis of bone tumors [12], computed tomography (CT) has unique advantages in spinal tumors, and can show the destruction pattern, surrounding invasion, and periosteal reaction pattern of bone lesions without structural overlap. In the era of computer-aided diagnosis, quantitative imaging techniques coupled with machine learning have facilitated the progression of data science research toward clinical translation [13]. Improved imaging methods yielding high-dimensional data enable a more detailed assessment of the macroscopic, cellular, and molecular properties of tumors [14]. Currently, there has been no meaningful research aimed at a non-invasive, comprehensive, and reproducible assessment of RANKL expression levels in spinal GCTB to facilitate predictive decision-making support.

In this study, we evaluate the prognostic role of RANKL status in patients with spinal GCTB, based on postoperative progression-free survival (PFS). In addition, we establish a machine learning model to facilitate the predictive value of RANKL expression levels in spinal GCTB patients. This study demonstrates that the evaluation of RANKL status may support the personalized design of immunotherapy interventions in GCTB.

2. Materials and Methods

2.1. Patients

This study was approved by the Institutional Research Ethics Board, and adheres to the tenets prescribed by the Declaration of Helsinki (institutional review board "M2019460", National Clinical Trial number "NCT04952818"). Signed informed consent was waived. We identified 135 consecutive patients with spinal GCTB who underwent surgery in our institution from January 2009 to February 2022. Inclusion criteria were as follows: (a) patients who had a histopathological diagnosis of spinal GCTB; (b) baseline CT performed before surgery, and (c) patients treated by surgical resection.

Among these 135 patients, 28 patients were excluded from this study for the following reasons: (a) The CT scan was performed more than 2 weeks before surgery; (b) Patient had a previous history of biopsy or treatment for the tumor lesion; (c) poor image quality, e.g., with peri-lesion artifacts; (d) missed postoperative paraffin specimen for reanalysis; (e) cases with negative H3F3A (H3 histone, family 3A); and (f) incomplete basic clinical data.

A total of 107 patients (mean age 32.94 $\pm$ 12.99 years; M:F = 45:62; age range: 18–71) were included in the final study cohort (Figure 1). Taking September 2019 as the cut-off point, they were divided into a training cohort (n = 82) and a validation cohort (n = 25).

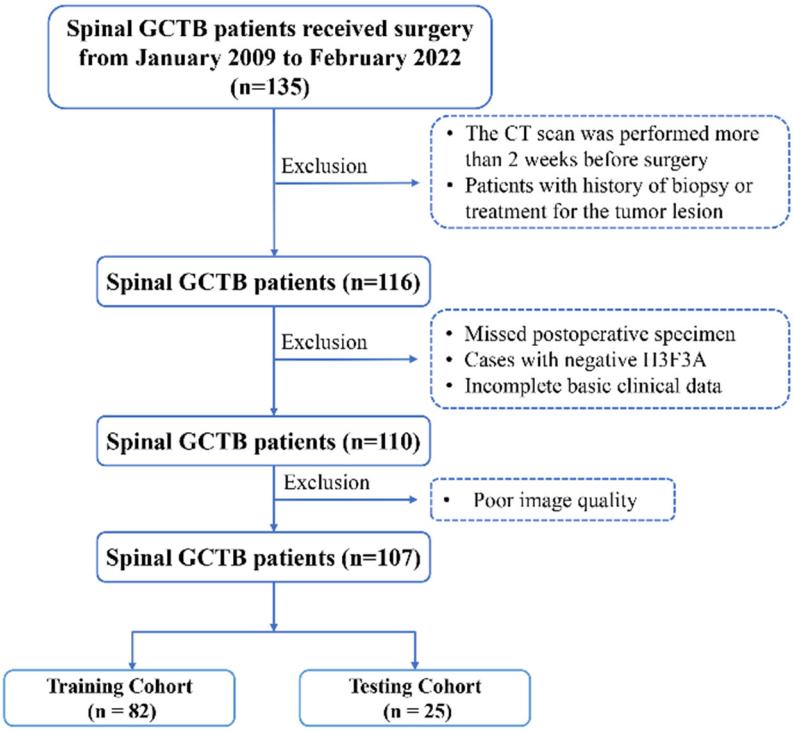


Figure 1. Study flow chart of patient selection. GCTB, giant cell tumor of bone; H3F3A, H3 histone family 3A.

2.2. Immunohistochemistry (IHC) Analysis

Samples from all cases were stained with hematoxylin-eosin (HE) and H3F3A (RevMAb Biosciences, South San Francisco, CA, USA, Catalog No.31-1145-00). IHC staining was performed using an anti-RANKL antibody from Abcam (No. ab222215). From each section, the degree of tumor positivity and the percentage of positive cells were evaluated by two independent pathologists who were blinded to clinical and pathological information. Similar to previous studies [3], RANKL staining was scored based on intensity (0 = no expression; 1 = weak to moderate expression; 2 = strong expression) and percentage of positive cells ($0 \leq 10\%$; 1 = 10–25%; 2 = 26–50%; $3 \geq 50\%$). Cutoff levels for the sum of scores were defined as 0–3 for low expression, and 4–5 for high expression.

2.3. Postoperative Follow-Up and Clinical Data

Patients were followed up according to the following guidelines: once every 3 months during the first 2 years; once every 6 months from 3–5 years; and annually after 5 years. Local disease recurrence was confirmed by MRI with an emerging mass found at the resection site with or without a tissue biopsy.

All the clinical data including age, gender, location, tumor stage (Enneking stage), radiotherapy, and denosumab treatment were obtained from the medical records. Simultaneously, imaging assessment of tumor size includes physician measurement (Y) of the longest diameter, while tumor volume is automatically measured by a manually constructed 3D tumor mask.

Based on previous studies, surgery and preoperative radiotherapy may have an impact on postoperative recurrence, so we included 75 patients who underwent total en bloc spondylectomy (TES) without preoperative radiotherapy for survival analysis to evaluate the prognostic significance of RANKL levels. Two cases were shown in Figures 2 and 3.

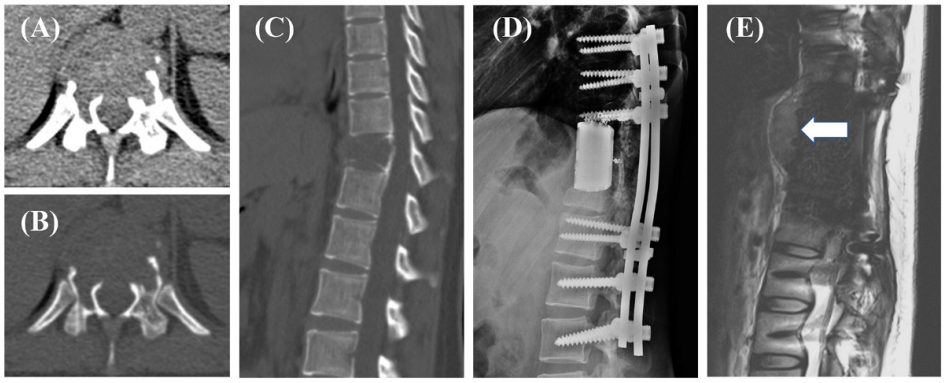


Figure 2. A 29-year-old female, axial (**A**,**B**) and sagittal (**C**) CT images showed the bone destruction of the T12, managed with total en bloc spondylectomy (TES, (**D**)). Postoperative pathology suggested a high expression of RANKL in the tumor. The patient did not undergo radiotherapy. At a 10-month follow-up review, recurrence was detected ((**E**), arrow).

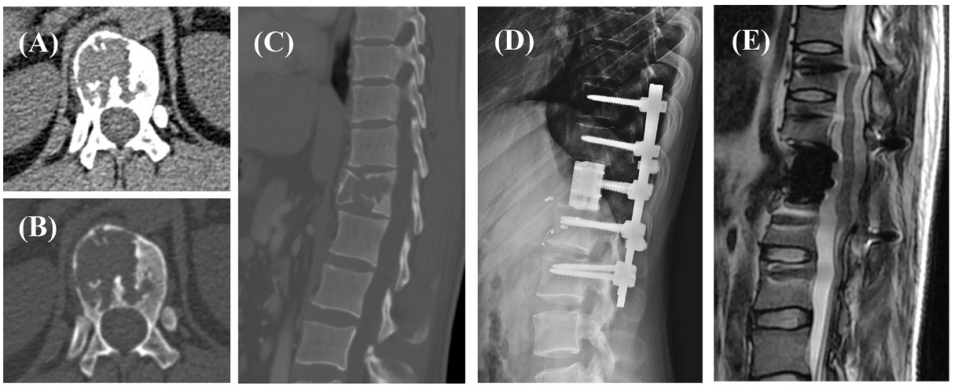


Figure 3. A 32-year-old male, CT images (**A**–**C**) showed a pathologic fracture of the T12 vertebra, managed with TES (**D**). Postoperative pathology showed low expression of RANKL without radiotherapy. At the 32-month follow-up, there was no evidence of recurrence (**E**).

2.4. Computed Tomography Imaging and Multiregional Labeling

CT scans were performed using a Sensation-64 scanner (SOMATOM Definition; Siemens, Erlangen, Germany) and a 64-slice spiral CT scanner (Light speed; GE Medical System, London, UK), with parameters including 120 kVp and automatic mAs. The collimator width was 0.625 and 0.60 mm, and the pitch was 1.0. All the CT images were retrieved from the picture archiving and communication system (PACS) in Digital Imaging and Communications in Medicine (DICOM) format. CT scans are isovolumetric and all images can be reconstructed for analysis.

Three different heterogeneous regions (VOI^{entire}, VOI^{edge}, and VOI^{core}) were segmented and labeled. Manual segmentation framework construction was conducted using Research Portal V1.1 (United Imaging Intelligence, Co., Ltd., Shanghai, China). The VOI^{entire} was segmented on axial CT images by a radiologist with 5 years of experience in spinal tumor diagnoses. Moreover, 35 patients were randomly selected and segmented by another musculoskeletal radiologist with 13 years of experience to construct a test-retest set and calculate the interclass correlation coefficients (ICCs) of radiomic features.

Next, VOI^{entire} was shrunk by 2 mm or 3 mm to generate the other two VOIs, the marginal area represents the transition zone of the bone tumor (VOI^{edge}), and the remaining core area represents the more central area of the bone tumor (VOI^{core}). In detail, VOI shrinkage was performed using the erosion function of the Research Portal V1.1. Figure 4 presents the schema for segmentation, radiomic feature extraction, and predictive modeling. It should be noted that we evaluated the selection of the width of the edge band (VOI^{edge}), and the detailed research process is shown in Supplementary Figure S1.

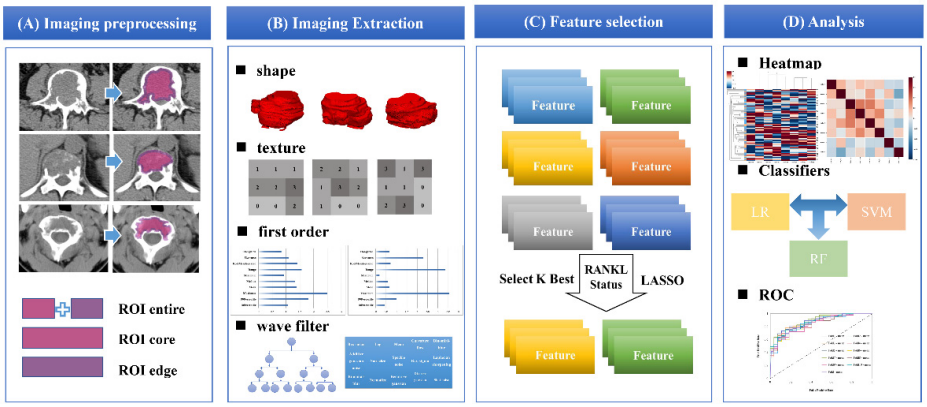


Figure 4. The workflow of necessary steps in this study. (**A**) The VOIs of tumors were manually segmented in axial CT images. Based on the entire tumor VOI, the VOIs of the edge zone and core area were automatically generated. (**B**) The radiomics features were extracted and processed to generate radiomics. (**C**) The sequential feature selection was performed by using Selected K Best and LASSO. (**D**) The RF, LR, and SVM were applied to build the classification models. The performance of the predictive models was evaluated with the area under the ROC curve. VOI, the volume of interest; RF, random forests; LR, logistic regression; SVM, support vector machines; ROC, receiver operating characteristic curve.

2.5. Radiomics Feature Extraction and Preprocessing

Radiomics feature selection, feature extraction, and machine learning models were established on the Research Portal V1.1 (United Imaging Intelligence, Co., Ltd., Shanghai, China). Prior to radiomic feature extraction, B-Spline interpolation resampling was used to normalize voxel size, and anisotropic voxels were resampled to generate isotropic voxels of 1.0 × 1.0 × 1.0 (mm). A total of 2264 radiomics features of each tumor lesion were extracted from preoperative CT imaging, including first-order features, shape features, texture features, and high-level features. The shape features were extracted according to the VOIs in the original image, and the remaining features were extracted in the original image and the filtered image.

Intraclass correlation coefficients (ICCs) were determined to assess the reproducibility of the radiomics features extracted from the VOIs drawn by three independent radiologists. Features with an ICC $\geq$ 0.8 were considered to be reliable. Before feature selection, all features were normalized by replacing outliers with the median of the particular variance vector and standardizing the data using the Z-score standardization.

2.6. Feature Selection and Machine Learning Analysis Strategy

Radiomics features were extracted from different VOIs (multiregions), and then Select K Best analysis and least absolute shrinkage and selection operator (LASSO) methods were used to analyze and screen the extracted image features related to RANKL expression. LASSO regularization involved a parameter, λ, to control the number of selected features; wherein a larger λ retains more features, and the final feature number was determined by λ to maximize the C-index in the training set. A multiple-feature-based radiomics signature (radiomics score, rad-score), was calculated for each patient using a linear combination of features that were each weighted by their respective coefficients.

Three classifiers were used in the prediction models: random forest (RF), support vector machine (SVM), and logistic regression (LR). According to events per independent variable (EPV) values in the multivariable prediction models [15], we selected the 3 features (10 EPV) with the highest efficiency to construct models from the three different heterogeneous regions (VOI^{entire}, VOI^{edge}, and VOI^{core}). At the same time, we compared the model performance of different edge bandwidths of 2 mm and 3 mm. Finally, 15 models were constructed and compared.

2.7. Statistical Analysis

Regarding continuous variables, data are expressed as mean ± SD or median interquartile range (IQR). For categorical variables, data are expressed as counts and percentages (*n*, %). The Wilcoxon signed-rank or Kruskal-Wallis tests were used to compare numerical variables, and Fisher's exact test was used to compare categorical variables. PFS probabilities were estimated using the Kaplan-Meier method and Cox proportional-hazards regression. Statistical analyses were conducted using SPSS (version 24.0, Chicago, IL, USA) and MedCalc (version 15.0, Mariakerke, Belgium). To assess whether the radiomic signature score could separate patients into low or high RANKL expression groups, the area under the curve (AUC) of the receiver operator characteristic (ROC) and its confidence interval (CI) were determined in accordance with the DeLong method. A result was considered statistically significant with a two-tailed *p*-value of $p < 0.05$.

3. Results

3.1. Patient Information

Of the 107 patients included in the final study cohort (Table 1), sixty-six (62%) spinal GCTB patients exhibited high levels of RANKL expression. The results and grading of immunohistochemical staining are shown in Figure 5A. In the follow-up cohort, RANKL levels were used to indicate a statistically significant difference in PFS, based on Kaplan-Meier survival analysis (Figure 5B). There was no difference in postoperative survival status based on patient age ($p = 0.190$) or gender ($p = 0.160$) (Supplementary Figure S2). No correlation was observed between patient age or gender and RANKL expression ($p > 0.05$), and these factors were not included in the machine learning analysis.

Table 1. Demographic and clinical information for the study cohort.

Variables	**Whole Cohort**		
RANKL status	Total	High	Low
	107	66	41
Gender			
Female	62	41	21
Male	45	25	20
Age			
Mean ± SD	32.94 ± 12.99	34.92 ± 13.22	29.76 ± 12.26
Location			
Cervical vertebrae	36	19	17
Thoracic vertebrae	35	18	17
Lumbar vertebrae	26	21	5
Sacrum	10	8	2
Tumor stage			
Stage II	40	26	14
Stage III	57	40	27
Tumor diameter, cm	4.72 ± 1.99	4.96 ± 1.94	4.34 ± 2.03
Tumor volume, cm^3	60.75 ± 147.72	54.76 ± 82.81	70.40 ± 215.71
Surgery method			
Partial removal/curettage	32	20	12
Total en bloc spondylectomy	75	45	30
Radiotherapy	22	13	9
Denosumab	15	8	7

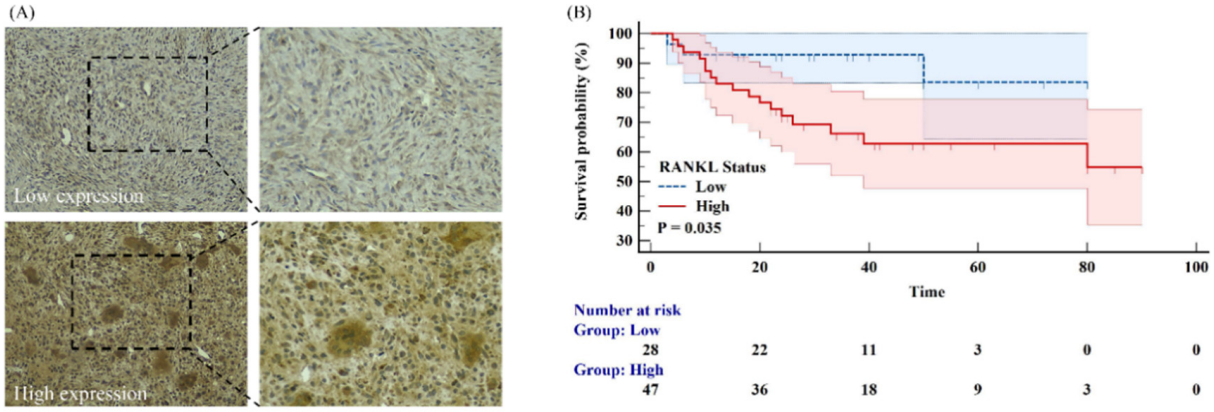


Figure 5. (**A**) Immunohistochemical staining of RANKL in spinal GCTB. Original amplification: 200×. (**B**) Kaplan-Meier survival analysis of the impact of RANKL status.

3.2. Feature Robustness for Multiregional VOIs

Quantification of all extracted features from each VOI is shown in Supplementary Table S1. The features extracted from the three different VOIs have different numbers of remaining valid features after being screened by the ICC calculation (threshold = 0.8). In the entire tumor vs. margin shrinkage segmentation, the stable feature rates were 51.06% (n = 1156), 43.95% (n = 995) and 54.37% (n = 1231) for VOI^{entire}, VOI^{edge}, and VOI^{core}, respectively. The number of stable features derived from 3D margin shrinkage segmentation was higher compared to the entire tumor and marginal zone (p = 0.038). The other groups were not significantly different, with p-values of p = 0.101 and p = 0.200, respectively. Table S1 details the number and percentage of stable features that were obtained with different 3D segmentation methods, grouped according to feature class. In the training cohort, with 10 repetitions of 10-fold cross-validation by the LASSO method, the top 3 features with the most significant coefficients were selected per repetition. Therefore, 300 features with different repetitions were extracted for each VOI. Features and repetitions are shown in Supplementary Figure S3 (edge band width 3 mm). The first three stable and effective features extracted based on each regional VOI were selected for the model construction.

3.3. Performance of Models Based on Different Classifiers

We compared the performance of models based on different VOIs and classifiers (Figure 6). For the radiomics models, the VOI^{core} model was the best performer; the AUC values of the three models are visually displayed in a heatmap diagram (Figure 7). For each VOI model, the performance of the RF classifier was better than that of SVM and LR, and the statistical indicators of DeLong's test are given in Supplementary Table S2. When the width of the edge band is 3 mm, the model performance of both the core region VOI and the edge region VOI is improved. The corresponding performance parameters of each model (edge width 3 mm), including AUC, F1 score, recall, precision, sensitivity, specificity, and accuracy, are given in Table 2.

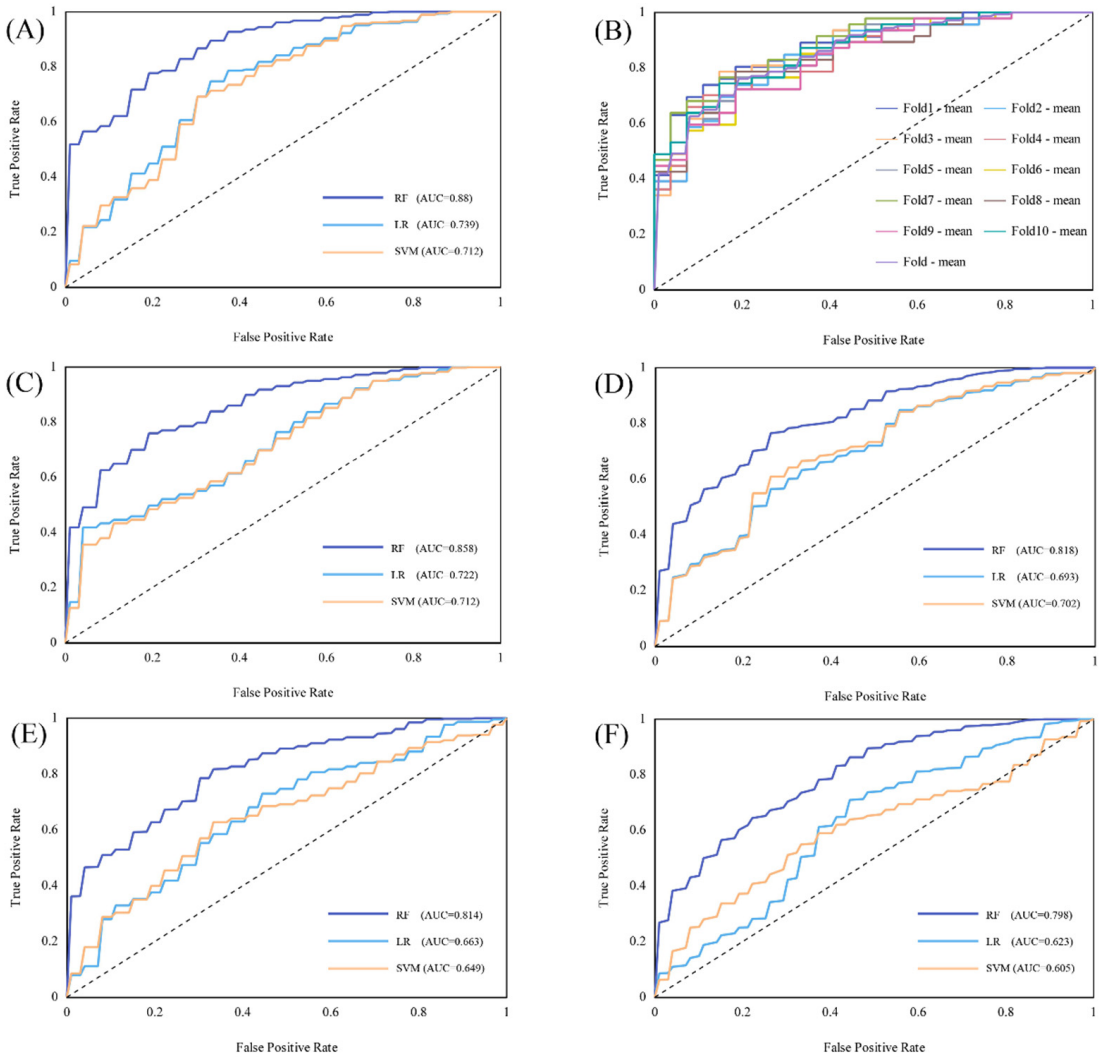


Figure 6. Receiver operating characteristic (ROC) curve demonstrating the performance radiomics models constructed by different classifiers (RF, LR, and SVM) in predicting RANKL status in spinal GCTB based on multiregional VOIs. (**A**) Entire tumor VOI (mean of 10-fold cross-validation); (**B**) Entire tumor VOI (each fold of 10-fold cross-validation); (**C**) Core area VOI (3 mm); (**D**) Edge zone VOI (3 mm); (**E**) Core area VOI (2 mm); (**F**) Edge zone VOI (2 mm). RF, random forests; LR, logistic regression; SVM, support vector machines.

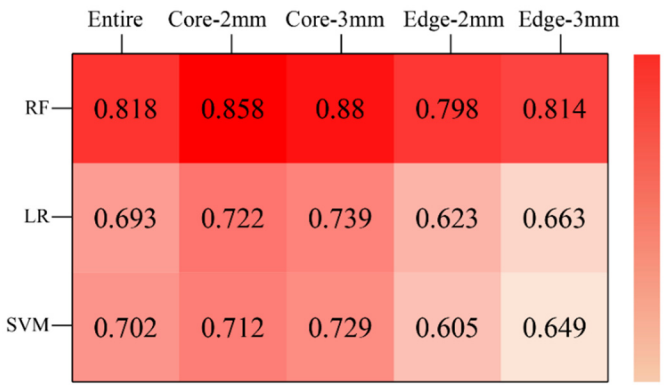


Figure 7. The heatmap of AUCs of three classifiers constructed with five different VOIs. Darker colors mean higher AUC values. AUC, area under the curve; RF, random forests; LR, logistic regression; SVM, support vector machines.

Table 2. The performance of the RF classifier is based on multiregional VOIs to distinguish RANKL states in the training cohort and validation cohort.

Index	AUC (95%CI)	F1 Score	Recall	Precision	Sensitivity	Specificity	Accuracy
Training cohort							
$RF_{VOI\text{-}entire}$	0.814 (0.718–0.912)	0.764	0.720	0.793	0.918	0.581	0.719
$RF_{VOI\text{-}edge}$	0.798 (0.695–0.902)	0.810	0.919	0.691	0.979	0.241	0.709
$RF_{VOI\text{-}core}$	0.880 (0.807–0.958)	0.857	0.934	0.817	0.720	0.719	0.802
Validation cohort							
$RF_{VOI\text{-}entire}$	0.648 (0.419–0.906)	0.674	0.657	0.708	0.657	0.433	0.646
$RF_{VOI\text{-}edge}$	0.677 (0.415–0.892)	0.790	0.940	0.688	0.940	0.233	0.682
$RF_{VOI\text{-}core}$	0.745 (0.571–0.898)	0.775	0.850	0.728	0.850	0.633	0.694

3.4. Performance and Validation of the Final Prediction Models

The RF model based on VOIs in the core region with 3 mm edge shrink was selected for independent validation. The features of different VOIs used to build the models are shown in Supplementary Table S3. The ROC curve and confusion matrix of the predicted results are shown in Figure 8. The columns represent the number of predicted classes, and the rows represent the number of true attribution classes of the data. The accuracy, precision, sensitivity, and specificity of the model are 0.640, 0.647, 0.785, and 0.454 respectively.

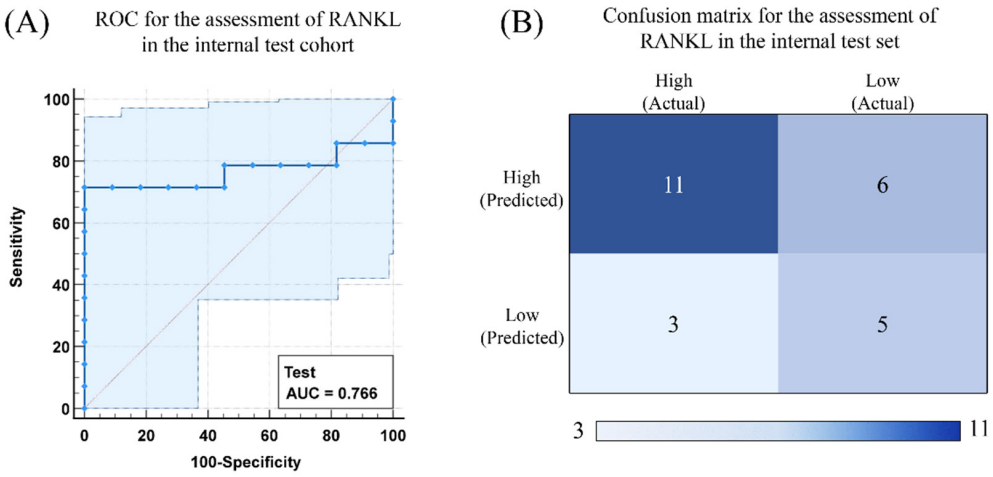


Figure 8. (**A**) ROC curves for prediction of RANKL status in the testing cohorts. (**B**) Confusion matrix for predicted outcomes of RANKL expression levels in the test cohort.

4. Discussion

In this study, we conducted a survival analysis to validate the prognostic value of the RANKL expression levels in a cohort of patients with a postoperative follow-up of spinal GCTB. At the same time, we applied machine learning methods based on radiomics signatures extracted from preoperative CT imaging to classify RANKL status in patients with spinal GCTB.

Treatment of GCTB cases located in the spine is challenging, as en bloc or wide resection is technically difficult and recurrence is common [6]. Improving the prognosis of GCTB of the spine is the focus of surgeons today. Considering the risks faced by patients undergoing the procedure, accurate preoperative stratification of patient prognosis can help in treatment plan planning and precise treatment. Using bioinformatics and combinatorial screening approaches to determine biomarker expression status could be useful in identifying patients who may benefit most from treatment. The RANK/RANKL pathway is often overexpressed and has been positively correlated with tumor progression and advanced disease in primary malignant tumors of the bone, including osteosarcoma, multiple myeloma, and GCTB [16–20]. Moreover, RANK and RANKL expressions are often higher in malignant histological subtypes of bone cancer. For example, RANKL

expression is often elevated in Stage III GCTB, and is a useful prognostic marker for predicting the risk of local disease recurrence [21]. Additionally, elevated RANK and RANKL may significantly increase the risk of metastasis [3]. In this study, we reveal for the first time that RANKL expression status is significantly correlated with disease prognosis in patients with spinal GCTB in a clinical cohort. This highlights that drugs targeting the RANK/RANKL pathway may effectively improve patient outcomes beyond merely inhibiting bone tissue destruction. Patients with recurrence with high RANKL expression, as in Figure 2, may be able to improve their prognosis if they are evaluated correctly early preoperatively, are able to suggest drug availability, surgeons expand resection, and postoperative radiation therapy.

Imaging is important in evaluating cancers for surgical planning, prognosis prediction, and post-treatment assessment. Advances in quantitative image analysis methods may offer a more comprehensive approach that includes spatiotemporal information, to produce image-driven biomarkers that may provide a deeper understanding of cancer biology, ultimately aiding in better clinical decisions [22,23]. Campanacci et al. applied conventional lesion size assessment and the new degree of CT ossification in a comparative analysis of 36 patients before and after denosumab treatment. In our cohort, some of the patients treated with denosumab (n = 15) partially undergo CT and/or MR scans. It is important to note that despite FDA approval in 2013, denosumab (Xgeva) has been available in our study countries since 2020. Among patients who had CT scans before and after treatment (n = 10), only 2 patients had local recurrence, implying that it may be difficult to statistically validate the predictive effect of the new CT evaluation proposed in this study in our cohort. Recurrence is a clinical observation, but it is worth mentioning that this article suggests that tumor treatment response is associated with postoperative recurrence, whereas in terms of microscopic expression, tumor treatment response is associated with therapeutic target status.

Currently, DNA sequencing and immunohistochemistry are the most accurate methods for molecular biomarker assessment in tumor. However, these methods have some limitations such as invasiveness, sampling error, and complications. Tumor multi-omics can be combined with predictive machine learning models, which could be the new digital method on the road to precision cancer medicine. In previous studies of biomarkers for tumor types such as lung cancer, glioma, breast cancer, and prostate cancer, radiomics is found to have the potential as a means to non-invasively predict the status of tumor biomarkers [24–31]. Radiomics approaches combined with a noninvasive machine learning model with tumor immunohistochemistry could improve treatment selection. Our study developed a radiomic signature of RANKL expression status in spinal GCTB based on CT imaging. The optimal model had an AUC of 0.88 in the training cohort, 0.745 in the validation cohort, and 0.766 in the test set. This radiomic signature could assist surgeons in evaluating patient prognosis before surgery and predicting drug treatment response, which would be helpful for clinical decision-making. Previous findings have also applied radiomics for preoperative differential diagnosis, recurrence prediction, and biomarker assessment of GCTB [32–34]. Although our study is preliminary, this may be a promising method for improving the precise treatment of spinal GCTB.

This study used three classifiers to construct models based on different VOIs to predict RANKL status in spinal GCTB patients. A total of 15 models were built and compared by ROC analysis. According to the AUCs, the RF classifier was better than the other two classifiers. Previous studies demonstrated the ideal classifiers for the construction of models may differ across organs, and RF may also inform prognostic models for other malignancies, including hepatocellular carcinoma [35], stomach cancer [36–38], rectal cancer [39,40], glioma [41], lung cancer [42], pelvic bone tumor [43], and soft tissue tumors [44]. The RF classifier also showed the best performance in recognizing the RANKL status of patients in this study, regardless of the type of VOI.

In radiomics studies, too many features can be selected that do not match the sample size in modeling, which may lead to overfitting. In our study, we follow the 10 EPV rule in binary classification problems [45–47]. According to the number of the smallest group in

our study cohort (n = 30), three features were chosen for the construction of the different models. At the same time, to ensure the robustness of the analysis, we selected the top three most important (in terms of efficacy and frequency) radiomics features with 100 repetitions. This may be a way to improve the repeatability and reproducibility of radiomics. We also summarized the features extracted from multiregional VOIs to compare feature stability. The number of stable features derived from VOI^{core} was higher compared to VOI^{entire}, and the features from the VOI^{core} are more concentrated. The RF model based on the tumor core region (VOI^{core}) achieved the highest AUC. The study clarified that RF classifiers with the core region of the tumor might perform better in CT image-based GCTB-related tasks.

Some limitations of our study are worth noting. First, this was a single-institution study with a limited sample size, a common concern in "data-hungry" AI-based research. Although our institution is a regionally renowned hospital specializing in spinal tumor surgery, it is quite difficult to expand the sample size due to the incidence of GCTB in the spine. In the future, multicenter studies may improve the robustness and general feasibility of radiomics results. Second, our study did not discuss the effects of surgical procedures on postoperative survival, but only performed survival analysis in patients after TES. TES is the mainstream surgical method for spinal tumors at present, and striving for total vertebral resection is the main consideration of the surgeon unless it cannot be achieved. Third, enhanced CT and MRI images were not analyzed in the retrospective design. In our cohort, preoperative CT without contrast was the most commonly performed (n = 107), while relatively fewer patients underwent enhanced CT (n = 43) or MRI (n = 65). Machine learning attempts based on other scanning modalities may be promising in the future. Finally, due to the limited scope of this study, the generalizability of our findings needs to be confirmed on additional datasets.

5. Conclusions

RANKL expression status may be an important molecular marker for the postoperative survival of spinal GCTB patients. We present an RF-based machine learning method that exhibited excellent AUC and stability in assessing RANKL status in patients with spinal GCTB. Radiomics features based on 3D margin shrinkage segmentation showed good performance in the preoperative evaluation of low/high RANKL expression.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14215201/s1, Figure S1: Two patients with different lesion sizes; Figure S2: Kaplan-Meier survival analysis for gender and age; Figure S3: Feature names extracted from different regions and the number of repetitions. Table S1: The total extracted features of each VOI, the number and percentage of features after ICC inspection of VOIs in different regions; Table S2: Delong test results between different models. Table S3: Features selected by the multi-region VOI models.

Author Contributions: Conceptualization, Q.W. and N.L.; methodology, Y.C. and X.L.; software, X.L.; validation, K.L. and P.X.; formal analysis, Q.W., K.L. and W.Z.; investigation, Q.W., S.Q., W.Z. and N.L.; resources, H.Y. and N.L.; data curation, Q.W., Y.C., S.Q. and P.X.; writing—original draft preparation, Q.W.; writing—review and editing, Q.W., Y.C., H.Y. and N.L.; visualization, X.L.; supervision, N.L.; project administration, H.Y. and N.L.; funding acquisition, H.Y. and N.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (grant number 81971578), Beijing Natural Science Foundation (grant number Z190020) and National Natural Science Foundation of China (grant number 81871326).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Ethics Committee of Peking University Third Hospital (IRB protocol M2019460, December 2022).

Informed Consent Statement: Patient consent was waived due to retrospective study.

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful to Zhukai Cong and Ming Zhang for the pathological analysis, and Heng Wang for providing technology platform services. We would like to thank the native English speaking scientists of Elixigen Company (Huntington Beach, CA, USA) for editing our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Casimiro, S.; Vilhais, G.; Gomes, I.; Costa, L. The Roadmap of RANKL/RANK Pathway in Cancer. *Cells* **2021**, *10*, 1978. [CrossRef]
2. Dougall, W.C. Molecular pathways: Osteoclast-dependent and osteoclast-independent roles of the RANKL/RANK/OPG pathway in tumorigenesis and metastasis. *Clin. Cancer Res.* **2012**, *18*, 326–335. [CrossRef] [PubMed]
3. Quattrini, I.; Pollino, S.; Pazzaglia, L.; Conti, A.; Novello, C.; Ferrari, C.; Pignotti, E.; Picci, P.; Benassi, M.S. Prognostic role of nuclear factor/IB and bone remodeling proteins in metastatic giant cell tumor of bone: A retrospective study. *J. Orthop. Res. Off. Publ. Orthop. Res. Soc.* **2015**, *33*, 1205–1211. [CrossRef] [PubMed]
4. Chawla, S.; Blay, J.Y.; Rutkowski, P.; Le Cesne, A.; Reichardt, P.; Gelderblom, H.; Grimer, R.J.; Choy, E.; Skubitz, K.; Seeger, L.; et al. Denosumab in patients with giant-cell tumour of bone: A multicentre, open-label, phase 2 study. *Lancet Oncol.* **2019**, *20*, 1719–1729. [CrossRef]
5. Tsukamoto, S.; Mavrogenis, A.F.; Kido, A.; Errani, C. Current Concepts in the Treatment of Giant Cell Tumors of Bone. *Cancers* **2021**, *13*, 3647. [CrossRef]
6. Bukata, S.V.; Blay, J.Y.; Rutkowski, P.; Skubitz, K.; Henshaw, R.; Seeger, L.; Dai, T.; Jandial, D.; Chawla, S. Denosumab Treatment for Giant Cell Tumor of the Spine Including the Sacrum. *Spine* **2021**, *46*, 277–284. [CrossRef]
7. Chen, X.; Li, H.; Zhu, S.; Wang, Y.; Qian, W. Pre-operative denosumab is associated with higher risk of local recurrence in giant cell tumor of bone: A systematic review and meta-analysis. *BMC Musculoskelet. Disord.* **2020**, *21*, 256. [CrossRef]
8. Palmerini, E.; Chawla, N.S.; Ferrari, S.; Sudan, M.; Picci, P.; Marchesi, E.; Leopardi, M.P.; Syed, I.; Sankhala, K.K.; Parthasarathy, P.; et al. Denosumab in advanced/unresectable giant-cell tumour of bone (GCTB): For how long? *Eur. J. Cancer* **2017**, *76*, 118–124. [CrossRef]
9. Asano, N.; Saito, M.; Kobayashi, E.; Morii, T.; Kikuta, K.; Watanabe, I.; Anazawa, U.; Takeuchi, K.; Suzuki, Y.; Susa, M.; et al. Preoperative Denosumab Therapy Against Giant Cell Tumor of Bone is Associated with an Increased Risk of Local Recurrence After Curettage Surgery. *Ann. Surg. Oncol.* **2022**, *29*, 3992–4000. [CrossRef]
10. Allard, B.; Aspeslagh, S.; Garaud, S.; Dupont, F.A.; Solinas, C.; Kok, M.; Routy, B.; Sotiriou, C.; Stagg, J.; Buisseret, L. Immuno-oncology-101: Overview of major concepts and translational perspectives. *Semin. Cancer Biol.* **2018**, *52*, 1–11. [CrossRef] [PubMed]
11. Kang, C.Y.; Duarte, S.E.; Kim, H.S.; Kim, E.; Park, J.; Lee, A.D.; Kim, Y.; Kim, L.; Cho, S.; Oh, Y.; et al. Artificial Intelligence-based Radiomics in the Era of Immuno-oncology. *Oncologist* **2022**, *27*, e471–e483. [CrossRef] [PubMed]
12. Von Schacky, C.E.; Wilhelm, N.J.; Schafer, V.S.; Leonhardt, Y.; Gassert, F.G.; Foreman, S.C.; Gassert, F.T.; Jung, M.; Jungmann, P.M.; Russe, M.F.; et al. Multitask Deep Learning for Segmentation and Classification of Primary Bone Tumors on Radiographs. *Radiology* **2021**, *301*, 398–406. [CrossRef] [PubMed]
13. Avanzo, M.; Wei, L.; Stancanello, J.; Vallieres, M.; Rao, A.; Morin, O.; Mattonen, S.A.; El Naqa, I. Machine and deep learning methods for radiomics. *Med. Phys.* **2020**, *47*, e185–e202. [CrossRef]
14. Limkin, E.J.; Sun, R.; Dercle, L.; Zacharaki, E.I.; Robert, C.; Reuze, S.; Schernberg, A.; Paragios, N.; Deutsch, E.; Ferte, C. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann. Oncol.* **2017**, *28*, 1191–1206. [CrossRef]
15. Peduzzi, P.; Concato, J.; Feinstein, A.R.; Holford, T.R. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* **1995**, *48*, 1503–1510. [CrossRef]
16. De Groot, A.F.; Appelman-Dijkstra, N.M.; van der Burg, S.H.; Kroep, J.R. The anti-tumor effect of RANKL inhibition in malignant solid tumors—A systematic review. *Cancer Treat. Rev.* **2018**, *62*, 18–28. [CrossRef] [PubMed]
17. Renema, N.; Navet, B.; Heymann, M.F.; Lezot, F.; Heymann, D. RANK-RANKL signalling in cancer. *Biosci. Rep.* **2016**, *36*, e00366. [CrossRef]
18. Gonzalez-Suarez, E.; Sanz-Moreno, A. RANK as a therapeutic target in cancer. *Febs J.* **2016**, *283*, 2018–2033. [CrossRef] [PubMed]
19. McGrath, E.E. OPG/RANKL/RANK Pathway as a Therapeutic Target in Cancer. *J. Thorac. Oncol.* **2011**, *6*, 1468–1473. [CrossRef] [PubMed]
20. Van Dam, P.A.; Verhoeven, Y.; Trinh, X.B.; Wouters, A.; Lardon, F.; Prenen, H.; Smits, E.; Baldewijns, M.; Lammens, M. RANK/RANKL signaling inhibition may improve the effectiveness of checkpoint blockade in cancer treatment. *Crit. Rev. Oncol. Hemat.* **2019**, *133*, 85–91. [CrossRef] [PubMed]
21. Ghani, S.A.; Wan Ismail, W.F.; Md Salleh, M.S.; Yahaya, S.; Syahrul Fitri, Z.M. The Values of Receptor Activator Nuclear Kappa-B Ligand Expression in Stage III Giant Cell Tumor of the Bone. *Indian J. Orthop.* **2018**, *52*, 31–34. [CrossRef] [PubMed]
22. El Naqa, I.; Ten Haken, R.K. Can radiomics personalise immunotherapy? *Lancet Oncol.* **2018**, *19*, 1138–1139. [CrossRef]
23. Gillies, R.J.; Anderson, A.R.; Gatenby, R.A.; Morse, D.L. The biology underlying molecular imaging in oncology: From genome to anatome and back again. *Clin. Radiol.* **2010**, *65*, 517–521. [CrossRef]

24. Calabrese, E.; Rudie, J.D.; Rauschecker, A.M.; Villanueva-Meyer, J.E.; Clarke, J.L.; Solomon, D.A.; Cha, S. Combining radiomics and deep convolutional neural network features from preoperative MRI for predicting clinically relevant genetic biomarkers in glioblastoma. *Neurooncol. Adv.* **2022**, *4*, vdac060. [CrossRef] [PubMed]
25. Chen, H.; Lin, F.; Zhang, J.; Lv, X.; Zhou, J.; Li, Z.C.; Chen, Y. Deep Learning Radiomics to Predict PTEN Mutation Status From Magnetic Resonance Imaging in Patients With Glioma. *Front. Oncol.* **2021**, *11*, 734433. [CrossRef]
26. Zhang, S.; Wang, X.; Yang, Z.; Zhu, Y.; Zhao, N.; Li, Y.; He, J.; Sun, H.; Xie, Z. Intra- and Peritumoral Radiomics Model Based on Early DCE-MRI for Preoperative Prediction of Molecular Subtypes in Invasive Ductal Breast Carcinoma: A Multitask Machine Learning Study. *Front. Oncol.* **2022**, *12*, 905551. [CrossRef]
27. D'Arnese, E.; Donato, G.W.D.; Sozzo, E.D.; Sollini, M.; Sciuto, D.; Santambrogio, M.D. On the Automation of Radiomics-Based Identification and Characterization of NSCLC. *IEEE J. Biomed. Health Inf.* **2022**, *26*, 2670–2679. [CrossRef]
28. Bao, J.; Liu, Y.; Ping, X.; Zha, X.; Hu, S.; Hu, C. Preoperative Ki-67 proliferation index prediction with a radiomics nomogram in stage T1a-b lung adenocarcinoma. *Eur. J. Radiol.* **2022**, *155*, 110437. [CrossRef]
29. Cui, R.; Yang, Z.; Liu, L. What does radiomics do in PD-L1 blockade therapy of NSCLC patients? *Thorac. Cancer* **2022**, *13*, 2669–2680. [CrossRef]
30. Liu, Y.F.; Shu, X.; Qiao, X.F.; Ai, G.Y.; Liu, L.; Liao, J.; Qian, S.; He, X.J. Radiomics-Based Machine Learning Models for Predicting P504s/P63 Immunohistochemical Expression: A Noninvasive Diagnostic Tool for Prostate Cancer. *Front. Oncol.* **2022**, *12*, 911426. [CrossRef]
31. Wang, J.; Hu, Y.; Zhou, X.; Bao, S.; Chen, Y.; Ge, M.; Jia, Z. A radiomics model based on DCE-MRI and DWI may improve the prediction of estimating IDH1 mutation and angiogenesis in gliomas. *Eur. J. Radiol.* **2022**, *147*, 110141. [CrossRef] [PubMed]
32. Yin, P.; Mao, N.; Zhao, C.; Wu, J.; Sun, C.; Chen, L.; Hong, N. Comparison of radiomics machine-learning classifiers and feature selection for differentiation of sacral chordoma and sacral giant cell tumour based on 3D computed tomography features. *Eur. Radiol.* **2019**, *29*, 1841–1847. [CrossRef] [PubMed]
33. Wang, Q.; Zhang, Y.; Zhang, E.; Xing, X.; Chen, Y.; Nie, K.; Yuan, H.; Su, M.Y.; Lang, N. A Multiparametric Method Based on Clinical and CT-Based Radiomics to Predict the Expression of p53 and VEGF in Patients With Spinal Giant Cell Tumor of Bone. *Front. Oncol.* **2022**, *12*, 894696. [CrossRef] [PubMed]
34. Wang, Q.; Zhang, Y.; Zhang, E.; Xing, X.; Chen, Y.; Su, M.Y.; Lang, N. Prediction of the early recurrence in spinal giant cell tumor of bone using radiomics of preoperative CT: Long-term outcome of 62 consecutive patients. *J. Bone Oncol.* **2021**, *27*, 100354. [CrossRef]
35. Chong, H.H.; Yang, L.; Sheng, R.F.; Yu, Y.L.; Wu, D.J.; Rao, S.X.; Yang, C.; Zeng, M.S. Multi-scale and multi-parametric radiomics of gadoxetate disodium-enhanced MRI predicts microvascular invasion and outcome in patients with solitary hepatocellular carcinoma $\leq$ 5 cm. *Eur. Radiol.* **2021**, *31*, 4824–4838. [CrossRef]
36. Wang, Y.; Liu, W.; Yu, Y.; Liu, J.J.; Xue, H.D.; Qi, Y.F.; Lei, J.; Yu, J.C.; Jin, Z.Y. CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer. *Eur. Radiol.* **2020**, *30*, 976–986. [CrossRef] [PubMed]
37. Wang, Y.; Liu, W.; Yu, Y.; Han, W.; Liu, J.J.; Xue, H.D.; Lei, J.; Jin, Z.Y.; Yu, J.C. Potential value of CT radiomics in the distinction of intestinal-type gastric adenocarcinomas. *Eur. Radiol.* **2020**, *30*, 2934–2944. [CrossRef]
38. Wang, Y.; Liu, W.; Yu, Y.; Liu, J.J.; Jiang, L.; Xue, H.D.; Lei, J.; Jin, Z.; Yu, J.C. Prediction of the Depth of Tumor Invasion in Gastric Cancer: Potential Role of CT Radiomics. *Acad. Radiol.* **2020**, *27*, 1077–1084. [CrossRef]
39. Horvat, N.; Veeraraghavan, H.; Khan, M.; Blazic, I.; Zheng, J.; Capanu, M.; Sala, E.; Garcia-Aguilar, J.; Gollub, M.J.; Petkovska, I. MR Imaging of Rectal Cancer: Radiomics Analysis to Assess Treatment Response after Neoadjuvant Therapy. *Radiology* **2018**, *287*, 833–843. [CrossRef]
40. Taghavi, M.; Trebeschi, S.; Simões, R.; Meek, D.B.; Beckers, R.C.J.; Lambregts, D.M.J.; Verhoef, C.; Houwers, J.B.; van der Heide, U.A.; Beets-Tan, R.G.H.; et al. Machine learning-based analysis of CT radiomics model for prediction of colorectal metachronous liver metastases. *Abdom. Radiol.* **2021**, *46*, 249–256. [CrossRef] [PubMed]
41. Kim, M.; Jung, S.Y.; Park, J.E.; Jo, Y.; Park, S.Y.; Nam, S.J.; Kim, J.H.; Kim, H.S. Diffusion- and perfusion-weighted MRI radiomics model may predict isocitrate dehydrogenase (IDH) mutation and tumor aggressiveness in diffuse lower grade glioma. *Eur. Radiol.* **2020**, *30*, 2142–2151. [CrossRef] [PubMed]
42. Jia, T.Y.; Xiong, J.F.; Li, X.Y.; Yu, W.; Xu, Z.Y.; Cai, X.W.; Ma, J.C.; Ren, Y.C.; Larsson, R.; Zhang, J.; et al. Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *Eur. Radiol.* **2019**, *29*, 4742–4750. [CrossRef] [PubMed]
43. Yin, P.; Zhi, X.; Sun, C.; Wang, S.; Liu, X.; Chen, L.; Hong, N. Radiomics Models for the Preoperative Prediction of Pelvic and Sacral Tumor Types: A Single-Center Retrospective Study of 795 Cases. *Front. Oncol.* **2021**, *11*, 709659. [CrossRef] [PubMed]
44. Fields, B.K.K.; Demirjian, N.L.; Hwang, D.H.; Varghese, B.A.; Cen, S.Y.; Lei, X.; Desai, B.; Duddalwar, V.; Matcuk, G.R., Jr. Whole-tumor 3D volumetric MRI-based radiomics approach for distinguishing between benign and malignant soft tissue tumors. *Eur. Radiol.* **2021**, *31*, 8522–8535. [CrossRef]
45. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **2016**, *278*, 563–577. [CrossRef]

46. Peduzzi, P.; Concato, J.; Kemper, E.; Holford, T.R.; Feinstein, A.R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **1996**, *49*, 1373–1379. [CrossRef]
47. Van Smeden, M.; Moons, K.G.; de Groot, J.A.; Collins, G.S.; Altman, D.G.; Eijkemans, M.J.; Reitsma, J.B. Sample size for binary logistic prediction models: Beyond events per variable criteria. *Stat. Methods Med. Res.* **2019**, *28*, 2455–2474. [CrossRef]