MRI-based delta-radiomics predicts pathologic complete response in high-grade soft-tissue sarcoma patients treated with neoadjuvant therapy

Radiotherapy and Oncology 164 (2021) 73–82

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

MRI-based delta-radiomics predicts pathologic complete response in high-grade soft-tissue sarcoma patients treated with neoadjuvant therapy

Jan C. Peeken [a,b,c,d,e,*], Rebecca Asadpour [a], Katja Specht [f], Eleanor Y. Chen [g], Olena Klymenko [a], Victor Akinkuoroye [a], Daniel S. Hippe [h], Matthew B Spraker [i], Stephanie K. Schaub [d], Hendrik Dapper [a], Carolin Knebel [j], Nina A. Mayr [d], Alexandra S. Gersing [k], Henry C. Woodruff [e,l], Philippe Lambin [e,l], Matthew J. Nyflot [d,m,1], Stephanie E. Combs [a,b,c,1]

[a] Department of Radiation Oncology, Klinikum rechts der Isar, Technical University of Munich (TUM); [b] Institute of Radiation Medicine (IRM), Department of Radiation Sciences (DRS), Helmholtz Zentrum, München; [c] Deutsches Konsortium für Translationale Krebsforschung (DKTK), Partner Site Munich, Germany; [d] Department of Radiation Oncology, University of Washington, Seattle, United States; [e] Department of Precision Medicine, GROW – School for Oncology and Developmental Biology, Maastricht University, The Netherlands; [f] Institute of Pathology, Technical University of Munich, Germany; [g] Department of Laboratory Medicine and Pathology, University of Washington; [h] Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle; [i] Department of Radiation Oncology, Washington University in St. Louis, United States; [j] Department of Orthopedics and Sports Orthopedics, Klinikum rechts der Isar; [k] Department of Radiology, Klinikum rechts der Isar, Technical University of Munich (TUM), Germany; [l] Department of Radiology and Nuclear Imaging, GROW- School for Oncology and Developmental Biology, Maastricht University Medical Centre, The Netherlands; [m] Department of Radiology, University of Washington, Seattle, United States

ARTICLE INFO

Article history:
Received 26 April 2021
Received in revised form 15 August 2021
Accepted 27 August 2021
Available online 4 October 2021

Keywords:
Soft-tissue sarcoma
Delta radiomics
Neoadjuvant radiotherapy
Machine learning
Response prediction
MRI

ABSTRACT

Purpose: In high-grade soft-tissue sarcomas (STS) the standard of care encompasses multimodal therapy regimens. While there is a growing body of evidence for prognostic pretreatment radiomic models, we hypothesized that temporal changes in radiomic features following neoadjuvant treatment ("delta-radiomics") may be able to predict the pathological complete response (pCR).

Methods: MRI scans (T1-weighted with fat-saturation and contrast-enhancement (T1FSGd) and T2-weighted with fat-saturation (T2FS)) of patients with STS of the extremities and trunk treated with neoadjuvant therapy were gathered from two independent institutions (training: 103, external testing: 53 patients). pCR was defined as <5% viable cells. After segmentation and preprocessing, 105 radiomic features were extracted. Delta-radiomic features were calculated by subtraction of features derived from MRI scans obtained before and after neoadjuvant therapy. After feature reduction, machine learning modeling was performed in 100 iterations of 3-fold nested cross-validation. Delta-radiomic models were compared with single timepoint models in the testing cohort.

Results: The combined delta-radiomic models achieved the best area under the receiver operating characteristic curve (AUC) of 0.75. Pre-therapeutic tumor volume was the best conventional predictor (AUC 0.70). The T2FS-based delta-radiomic model had the most balanced classification performance with a balanced accuracy of 0.69. Delta-radiomic models achieved better reproducibility than single timepoint radiomic models, RECIST or the peri-therapeutic volume change. Delta-radiomic models were significantly associated with survival in multivariate Cox regression.

Conclusion: This exploratory analysis demonstrated that MRI-based delta-radiomics improves prediction of pCR over tumor volume and RECIST. Delta-radiomics may one day function as a biomarker for personalized treatment adaptations.

The standard of care for high-grade soft-tissue sarcomas (STS) includes surgery, radiation therapy (RT) and/or chemotherapy (CTx). These treatment strategies achieve high local control rates but unfavorable overall survival (OS) and distant control [1-4].

RT can be delivered in the neoadjuvant or adjuvant setting. Compared to adjuvant RT, neoadjuvant RT offers several advantages including lower radiation doses, smaller target volumes, and reduced late toxicities [3,5]. The administration of chemotherapy remains more controversial. The phase-III EORTC62931 trial couldn't demonstrate a survival benefit for adjuvant chemotherapy

* Corresponding author at: Klinik und Poliklinik für RadioOnkologie und Strahlentherapie, Klinikum rechts der Isar, Technische Universität München (TUM), Ismaningestr. 22, 81675 München, Germany.
E-mail address: jan.peeken@tum.de (J.C. Peeken).
[1] These authors contributed equally.

https://doi.org/10.1016/j.radonc.2021.08.023
0167-8140/

[6]. The ISG-STS1001 trial, however, showed a survival benefit after neoadjuvant CTx compared to a "histology-tailored" treatment approach [7].

The benefit of neoadjuvant treatment concepts lays in the possibility to assess treatment response. This information could then be used for individual therapy escalation [8]. As a potential biomarker in STS patients, the pathological complete response (pCR) is currently being used in prospective trials as a surrogate marker for patients' outcomes [9]. A recent meta-analysis found a significant predictive value for OS [10].

Imaging constitutes an alternative tool to characterize tissue. As consequence, multiple authors have proposed quantitative imaging analysis ("radiomics") as a potential novel method to assess treatment response [11,12]. Radiomics is defined as a high-throughput quantitative analysis of imaging data [13]. Predefined features assessing the texture, intensity distribution, or shape of a volume of interest (VOI) are calculated and used as input for machine learning (ML) models [13-15]. Radiomics has been shown to predict various clinical and biological endpoints including pathological characteristics, prognosis, tumor progression, spatial infiltrations, and molecular aberrations in multiple cancer types [16-22]. In STS patients, there is growing evidence that radiomics can be successfully applied to predict OS, distant progression, and tumor grading using pretherapeutic imaging [23-29]. We hypothesize that temporal changes of radiomic features ("delta-radiomics") obtained before and after neoadjuvant therapy may be able to predict treatment response [30,31]. Crombé et al. first demonstrated prediction of pCR after neoadjuvant CTx using a delta-radiomics approach in STS patients in a monocentric study [32].

We analyzed the potential of MRI-based delta-radiomics to predict pCR in STS patients that received neoadjuvant RT and/or CTx. Radiomic analysis was performed using the two MRI sequences "contrast-enhanced and fat-saturated T1-weighted" (T1FSGd) and "fat-saturated T2-weighted" (T2FS) obtained before and after neoadjuvant therapies. The results were compared to tumor volume changes, Response Evaluation Criteria in Solid Tumours (RECIST), and single timepoint radiomic models. All models were externally validated.

Material & methods

Patients

Two independent patient cohorts were retrospectively collected at the University of Washington/Seattle Cancer Care Alliance, referred to as "training cohort", and the Technical University of Munich, referred to as "testing cohort". The inclusion criteria were patients with STS of the extremities and trunk treated with neoadjuvant RT with or without Ctx followed by tumor resection in curative intent. Exclusion criteria encompassed definitive, palliative, or postoperative RT, brachytherapy, other tumor locations, early abortion of RT (cut-off at 80% of planned total dose), osteosarcomas, Ewing sarcomas, rhabdomyosarcomas, endoprosthesis-dependent artifacts, missing pre- or post-therapeutic MRI scans, and incongruent image plane orientations between pre- and post-therapeutic MRI scans (see Supplemental Fig. 1 for a patient workflow). pCR values were obtained from all patients. If the information was missing, the surgical specimen was reassessed by board-certified pathologists at each institution (EC and KS) [33]. pCR was defined as less than 5% viable cells in the surgical specimen. The overall survival (OS) was calculated from initial pathologic diagnosis to the time point of death or the time point of censoring. Approval from the ethic committees was received (reference number 466/16 s). Informed consent was given before therapy. Data reporting follows the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) recommendations (Supplemental Material, Appendix) [34].

Image acquisition and segmentation

Each patient received pre-RT and post-RT MRI scans. Patients that received Ctx (neoadjuvant in all cases) had an additional MRI before Ctx administration. Table S1 describes acquisition parameters and scan planes. Tumor segmentation was performed manually (authors: MBS, JCP, VA; Supplemental Methods). To compensate for operator-dependent bias, multiple delineations were performed for 20 randomly selected patients on the pre-therapeutic MRI by three operators (authors: RA, MBS, JCP) in the training cohort (see Fig. 1). Dice similarity coefficients (DSC) were computed using 3D Slicer (DiceComputation module) [35].

Image preprocessing and radiomic feature extraction

See Supplemental Material for a detailed description of image preprocessing and radiomic feature extraction. In brief, preprocessing included bias field correction, intensity normalization, and isotropic resampling. 105 features per MR sequence were extracted using pyradiomics following the recommendations of the Imaging Biomarker Standardization Initiative (IBSI) [36,37]. Radiomic features included first-order, shape, and texture features (see Table S2 for a detailed listing). After ComBatHarmonization [38-42], delta-radiomic features were calculated by absolute subtraction of post-therapeutic feature values from pre-therapeutic feature values ($x_{pre} - x_{post}$). For single timepoint models, radiomic features from pre- and post-RT MRIs were directly used as model input. For an exploratory analysis, radiomic feature were also extracted from post-Ctx MRIs if available.

Feature reduction

All feature reduction steps were performed using the training cohort. First, all features susceptible to variations in segmentation performed in the subset of patients with multiple segmentations were excluded. As a threshold, an intraclass correlation coefficient (ICC) (3,1) of 0.8 was used. The remaining features (T2FS: 72, T1FSGd: 103) were then used as input for the modeling pipeline including additional feature reduction and model training (Fig. 1). The feature reduction procedure was performed using 1000 bootstrap samples. For each bootstrap sample, two feature selection stages were applied. First, highly intercorrelated features defined by a Spearman correlation coefficient of greater than 0.8 were excluded. For identified feature pairs, the feature with the highest mean correlation to all remaining features was excluded. Second, the Boruta algorithm was applied to filter the most relevant features [43,44]. The features were ranked according to the frequency of their selection in the 1000 bootstrap runs. The final feature set was defined as the n top-ranking features with n defined as the median feature number selected over all bootstrap runs.

Modeling strategy

For ML model comparison and unbiased performance evaluation on the training set, 100 iterations of three-fold nested cross-validation were performed based on the code by Deist et al. built upon the "caret" package [45,46]. Three common ML techniques were compared to predict pCR: random forest (RF), elastic net regression (ENR), and LogitBoost [47-49]. After Synthetic Minority Oversampling Technique (SMOTE) for imbalance correction (see Supplemental Methods), hyperparameters were optimized using

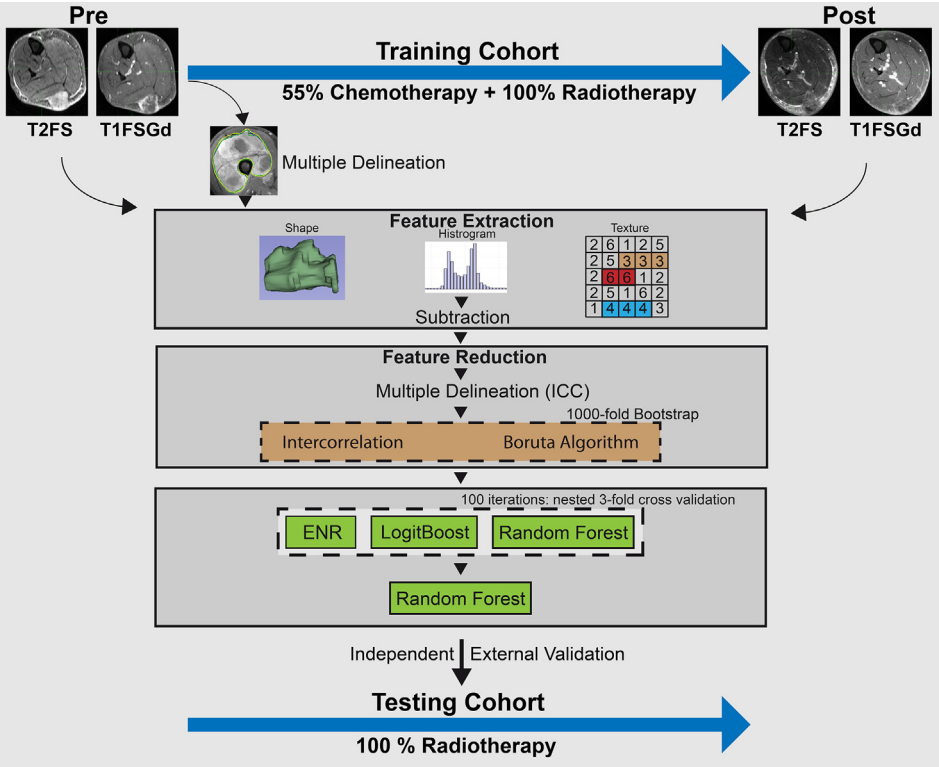

Fig. 1. The Delta-Radiomics Workflow. Abbreviations: Delta: delta-radiomics models, ENR: elastic net regression, ICC: intra-class correlation coefficient, pCR: pathological complete response, Pre: pre-radio(chemo)therapy models, Post: post-radio(chemo)therapy models, T1FSGd: contrast-enhanced T1-weighted fat saturated, T2FS: T2-weighted fat saturated.

grid search as part of the inner folds. See Table S3 for hyperparameter tuning spaces. The selected hyperparameters were then used for testing on the five outer folds. The mean receiver operating characteristic (ROC) curve (AUC) over all outer folds was calculated for model comparison. The hyperparameter combination with the best mean performance was used to retrain a final model on the whole training set. The final models were externally validated on the testing cohort. 95% confidence intervals (95% CI) were estimated using 1000-fold bootstrapping.

For all developed models the complete pipeline was applied separately. In total, three different delta-radiomic models were developed: "*Delta-T1FSGd*" based on T1FSGd, "*Delta-T2FS*" based on T2FS, and "*Delta-combined*" combining both feature sets. For comparison, single timepoint radiomic models were developed using the pre-therapeutic MRI scans: "*Pre-T1FSGd*", "*Pre-T2FS*", and "*Pre-combined*", or the post-therapeutic MRI scans: "*Post-T1FSGd*", "*Post-T2FS*", and "*Post-combined*". Moreover, the value of the AJCC staging system (8th edition) "*AJCC*", *RECIST* 1.1 (Supplemental Methods) [50], the tumor volume change "*Delta-Volume*", the pre-therapeutic tumor volume "*Pre-Volume*", and the post-therapeutic tumor volume "*Post-Volume*" were assessed. For quality assurance, the delta-radiomic models were retrained with permutated radiomic features. To assess the value of all models using a mixed cohort 100 iterations of three-fold nested cross-validation were applied as described above.

Exploratory analysis of survival and transferability to chemotherapy

Multivariate Cox proportional hazards regression was used to assess association with OS. The concordance index (C-index) was used to evaluate prognostic performance. Decision curve analysis was calculated for the multivariate models for OS at 3 years (see Supplemental Methods and caption of Fig. 3) [51,52].

We sought to evaluate if our developed models would function using delta-radiomic features calculated from MRIs obtained before and after neoadjuvant CTx, too (see Supplemental Fig. S2 for an overview of MRI timepoints). As in the testing cohort no patient received Ctx, our models were tested on the subset of the training cohort that received CTx. Chemotherapy-dependent delta-radiomic features we calculated by subtraction of radiomic feature calculated on the pre-therapeutic MRI and on the post-Ctx "interim" MRI ($x_{preCtx} - x_{postCtx}$).

Statistical analysis

Modeling and statistical analysis were performed using R (version 3.4.0, R core team, Vienna, Austria). Table S4 displays all R packages used. See Supplemental Methods for a description of outcome measures, feature importance, and calibration analysis. Delta-radiomic and multivariate models can be obtained from https://github.com/jacapan/Delta-Radiomics. Individual patient data may be available on reasonable request dependent on local ethics committee voting and in compliance with data protection rights.

Results

Patient demographics (training: 103, testing: 53 patients), staging groups, and RT doses were similar between the cohorts (Table 1). The frequency of neoadjuvant Ctx in addition to RT significantly differed between the training cohort 55% (n = 56) and the testing cohort 0% ($p < 0.001$). When CTx was administered, it was always preceded RT and 91% (n = 51) of patients received an AIM (anthracycline + ifosfamide + mesna)-based regimen (Table S5). Ctx was a significant prognostic factor for OS (p = 0.03) in the training cohort. There was a trend towards an

Table 1
Patient demographics, outcome and treatment specifics.

Institution		Training (UW)	Testing (TUM)	*p*-value[1]	Adjusted *p*-value
Accrual time		2008–2017	2010–2019		
Total Patients		102 p	59 p		
Patients	T1FSGd sequence	100 p (98%)	53 p (90%)		
	T2FS sequence	98 p (96%)	49 p (83%)		
	Both sequences	96 p (94%)	43 p (73%)		
Recurrent		0 p	5 p (8%)		
Age		m 54 (19–86)	m 57 (22–87)	0.331	1
Gender	female	38 (37%)	24 p (41%)	0.737	1
	male	64 (63%)	35 p (59%)		
T-stage[2]	1	6 p (6%)	1 p (2%)	0.432	1
	2	35 p (34%)	24 p (41%)		
	3	35 p (34%)	23 p (39%)		
	4	26 p (25%)	11 p (19%)		
N-stage[2]	0	102 p (100%)	58 p (98%)	0.366	1
	1	0 p (0%)	1 p (2%)		
Grading[3]	1	22 p (22%)	5 p (8%)	0.074	1
	2	39 p (38%)	23 p (39%)		
	3	41 p (40%)	31 p (57%)		
AJCC-Stage[2]	IA	2 p (2%)	0 p (0%)	0.16	1
	IB	20 p (20%)	5 p (8%)		
	II	4 p (4%)	1 p (2%)		
	IIIA	28 p (27%)	22 p (37%)		
	IIIB	45 p (44%)	31 p (53%)		
	IV	3 p (3%)	0 p (0%)		
Region	Upper Extremities	17 p (17%)	7 p (12%)	0.169	1
	Lower Extremities	75 p (74%)	51 p (87%)		
	Trunk	10 p (10%)	1 p (2%)		
Prognosis					
Median OS		Not reached	Not reached	0.8	1
Therapy information					
Margin-status	positive	20 p (20%)	7 p (12%)	0.026	0.38
	negative	81 p (80%)	48 p (81%)		
	x	0 p (0%)	4 p (7%)		
Total Dose		m 50 Gy (42–60 Gy)	m 50 Gy (50–56 Gy)	0.303	1
Chemotherapy		56 p (55%)	0 p (0%)	<0.001	<0.001
Viable cells after neoadjuvant therapy		m 40% (0–100%)	m 30% (0–100%)	0.150	1
pCR	positive	11 (11%)	5 (8%)	0.787	1
	negative	91 (89%)	54 (92%)		

Abbreviations: *: *p*-value < 0.05, pCR: pathological complete response, AJCC: American Joint Committee on Cancer staging system, m: median, p: patients, r: range, RT: radiation therapy.
[1] Wilcoxon rank-sum test for continuous and ordinal variables, Fisher's exact test for nominal variables, log-rank test for comparison of survival times. Corrected for multiple testing by Bonferroni correction ("*p*-value adjusted").
[2] Following AJCC staging manual 8th edition [61].
[3] According to the French Federation of Cancer Centers Sarcoma Group (FNCLCC).

uneven histology distribution (p = 0.087, Table S6). pCR was achieved in 11% (n = 11) and 8% (n = 5) of patients in the training and testing cohort, respectively (predominantly pleomorphic sarcoma and myxoid liposarcoma, see Table S7). The similarity between multiple VOI delineations was rated with a DSC of 0.91 (standard deviation 0.035).

In nested cross-validation, the RF models achieved the best mean performance with a mean AUC of 0.73 (AUC ENR: 0.62, AUC LB 0.67) (Table S8a). When ranking the ML models for each fold by their AUC value, RF, ENR, and LB achieved mean ranks of 1.4, 2.55, and 2.05, respectively (Table S8b). Thus, the RF algorithm was chosen to train the final models.

In the training set, the predictive performances in AUC were comparable for delta-radiomic models, *RECIST,* and *Delta-Volume* (AUCs: 0.74–0.80) (Fig. 2: AUC-values, Fig. 3: ROC-curves, Supplemental Fig. S3: Calibration-curves). In the testing cohort, *Delta-T1FSGd* and *Delta-combined* showed better performances that were more similar to the training cohort performances than for *Delta-T2FS* with AUC values of 0.70 (95% CI: 0.43–0.92, AUC difference: −0.05), 0.75 (95% CI: 0.56–0.93, AUC difference: −0.05), and 0.65 (95% CI: 0.33–0.92, AUC difference: −0.09), respectively. *RECIST* (AUC 0.6, 95% CI: 0.33–0.86), *Delta-Volume* (AUC 0.43, 95% CI: 0.06–0.82), and *AJCC* (AUC 0.51, 95% CI: 0.25–0.75) achieved predictive performances closer to random. After permutation of radiomic features no predictive performance above random was achieved in both cohorts (AUC-training: 0.38, 0.52, 0.59; AUC-testing: 0.45, 0.48, 0.47 for *Delta-T1FSGd*, *Delta-T2FS*, and *Delta-combined*, respectively)

Next, we evaluated the predictive performance of single timepoint models (Table 3: AUC-values, Supplemental Fig. S4: ROC curves). The T1FSGd-based models achieved better predictive performances than the T2FS-based models in both single timepoints in the training set. In the external validation cohort, however, the only single timepoint models with relevant predictive performance were *Post-combined* and *Pre-Volume* with AUC-values of 0.68 (95% CI: 0.44–0.91, AUC difference: −0.06) and 0.70 (95% CI: 0.48–0.9, AUC difference: +0.08).

Due to the imbalanced dataset, accuracy did not produce reliable metrics with high values for non-discriminative predictors. The best performances were achieved by *Delta-T2FS* (Matthew's correlation coefficient 0.48, balanced accuracy 0.69, F1-score: 0.5) (Table S9 displays all remaining models). Due to the imbalance

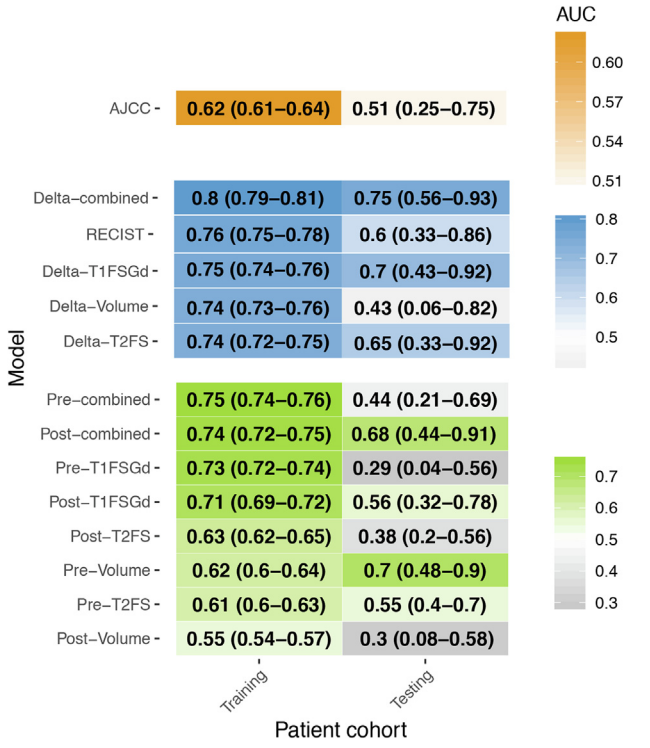

Fig. 2. Predictive performance of delta-radiomics, single timepoint radiomics, and clinical models. Area under the receiver operating characteristic curve (AUC) values for the prediction of pCR. 95 % confidence intervals are shown in parenthesis. Orange: clinical model, blue: delta-radiomics models/Volume change/RECIST, green: single timepoint models. Abbreviations: AJCC: American Joint Committee on Cancer staging system 8th Edition, AUC: area under the receiver operating characteristic curve, Delta: Delta-radiomics models, RECIST: Response Evaluation Criteria in Solid Tumors, T1FSGd: contrast-enhanced T1-weighted fat saturated, T2FS: T2-weighted fat saturated, pCR: pathological complete response, Pre: pre-radio(chemo)therapy, Post: post-radio(chemo)therapy models.

and small test set, calibration curves were "weak" for all models (Supplemental Fig. S3). The Hosmer-Lemeshow test was significant for all models reflecting the poor calibration of all models [53].

The final feature numbers used as input for modeling were 16, 15, and 31 features for *Delta-T1FSGd*, *Delta-T2FS*, and *Delta-combined*. Feature importance was assessed for the Delta-radiomics models (Table S10). For *Delta-T1FSGd,* the three most important features included two texture features "Busyness" and "SizeZoneNonUniformityNormalized", as well as the shape feature "Flatness". *Delta-T2FS* was also dominated by "Flatness" and two texture features "SizeZoneNonUniformity" and "ZoneEntropy". In the combined model, the important features from the single modality models also scored highest.

In an exploratory analysis we tested the association of the delta-radiomic models, *Pre-Volume*, and pCR with OS in the combined patient cohort corrected for *AJCC* and age as known prognostic variables. All delta-radiomic models, but not pCR and *Pre-Volume*, were significantly associated with OS in multivariate Cox regression (Table S11). In the training set, *Delta-T2FS*, *Delta-combined*, and *pCR* were significantly associated with OS, whereas in the test set none of the predictors, including age and AJCC, were associated with OS. The C-index of the clinical baseline model (AJCC+ age) of 0.68 in the training set could be improved by adding pCR (0.73), and *Delta-T2FS* (0.69) (Table S12). In the test set, however, only *Delta-combined* (0.69) and *Delta-T1FSGd* (0.71) improved the C-index of the clinical model (0.68). Likewise, both models achieved a net clinical benefit above the clinical model and the pCR-based multivariate model in decision curve analysis in the test set (Fig. 4). Exemplary patient cases are displayed in Fig. 5.

To determine the transferability to neoadjuvant chemotherapy delta-radiomics, we tested our delta-radiomic models using delta-radiomic features derived from MRIs obtained before and after neoadjuvant CTx on the subset of the training cohort that received CTx. *Delta-T1FSGd*, *Delta-T2FS*, and *Delta-combined* achieved AUC values of 0.90 (0.80–0.98), 0.88 (0.69–0.99), and 0.91 (0.74–1.00), respectively.

Because the two cohorts were found to have different patient characteristics, especially in use of chemotherapy and tumor histology, we performed a secondary analysis for all models using a mixed patient cohort. After 100 iterations of 3-fold nested cross-validation on the complete mixed cohort, the delta-radiomic models achieved the best performances (*Delta-combined*: AUC 0.79 (95% CI: 0.78–0.80), *Delta-T2FS* AUC 0.73 (95% CI: 0.72–0.75), *Delta-T1FSGd* 0.73 (95% CI: 0.72–0.74) compared to volume-based metrics (*RECIST* AUC 0.64 (95% CI: 0.63–0.66), *Delta-Volume* AUC 0.62 (95% CI: 0.6–0.64), *Pre-Volume* (0.66 (95% CI: 0.64–0.67)) or single time point radiomic models (best model: *Post-combined* AUC 0.64 (95% CI: 0.63–0.65)) (Supplemental Fig. S5). The delta-radiomics model performance was also comparable to the primary analysis (maximal testing AUC 0.75).

Discussion

We could demonstrate delta-radiomics-based response assessment in patients receiving neoadjuvant therapy. Delta-radiomic models achieved better reproducibility than single timepoint radiomic models, RECIST, or the peri-therapeutic volume change. The combined delta-radiomic model and T2FS-based model achieved the best predictive performance in terms of AUC. The best classification performance in terms of balanced accuracy was achieved by the T2FS-based model. Pre-therapeutic tumor volume was the best single timepoint predictor. While association with pCR was the primary hypothesis, the delta-radiomic models were significantly associated with OS in multivariate Cox regression models in a combined cohort.

While there is a growing body of evidence suggesting association between pre-treatment radiomics and outcomes in STS, data on delta-radiomics remains scarce. In a small 30-patient study, delta-radiomic analysis of diffusion-weighted MR imaging improved prediction of the treatment effect score (response threshold 50%) in internal cross validation [54]. Crombé et al. analyzed the value of T2-weighted sequence-based delta-radiomics for pCR predictions in STS patients after neoadjuvant Ctx in a monocentric cohort of 65 patients [55]. The final model achieved an AUC of 0.63 within a 15-patient holdout set. Lin et al. conducted a similar study in osteosarcoma patients using CT imaging. A performance with an AUC of 0.82 in an internal testing cohort was demonstrated [56]. Our delta-radiomic models had comparable performances with AUC-values ranging from 0.68 to 0.75. Since there were significant differences in histologies, imaging modalities, distribution of outcome variables, and treatment regimens, the performances within the above mentioned studies cannot be directly compared. One advantage of our study is that an independent test cohort was used providing a TRIPOD type III validation [34].

TRIPOD type III validation was achieved through the use of independent training and testing cohorts for our primary analysis [34]. However, as both cohorts were found to have significant differences in clinical features, histologies, and treatment regimens (e.g. chemotherapy), we performed a post-hoc secondary analysis with a mixed cohort and 3-fold nested cross validation to evaluate the effects of the cohort construction. Again, the delta-radiomic models outperformed volume-based metrics such as *RECIST* or single timepoint radiomic-models. Even though differences in the val-

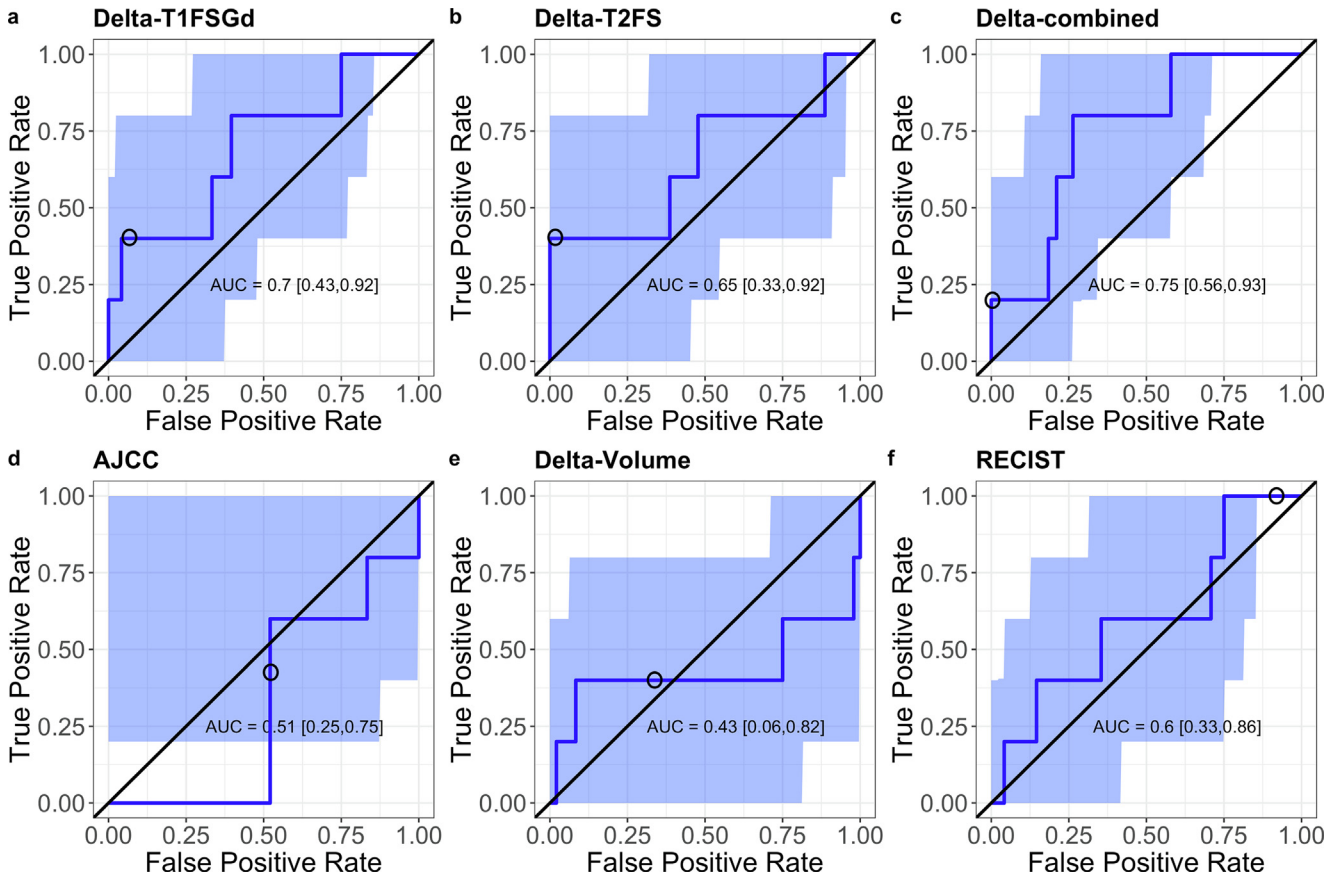

Fig. 3. Receiver operating characteristic curves (ROC) of the delta-radiomic models in the testing patient cohort. The shaded area represents the 95% confidence interval. The circles represent the cut-points applied for classification (median). Abbreviations: AJCC: American Joint Committee on Cancer staging system 8th edition, AUC: area under the ROC curve, Delta: delta-radiomics models, RECIST: Response Evaluation Criteria in Solid Tumors, T1FSGd: contrast-enhanced T1-weighted fat saturated, T2FS: T2-weighted fat saturated.

idation method mean the results cannot be directly compared, the performances in the mixed cohort were similar compared to that achieved on the primary training set.

In this study we used pCR as a surrogate marker for patient outcome instead of directly predicting OS, due to the small patient cohorts. pCR is often used in prospective trials as a surrogate marker reducing the need for long follow up and large patient numbers [9]. Ultimately, a response marker needs to prove its usefulness for patients' survival. We addressed this fact by performing exploratory multivariate Cox regression analyses. Interestingly, we could demonstrate an association of the delta-radiomic models with OS in the combined and training cohorts. pCR itself was only significant in the training cohort. With its low patient and event number, none of the prognostic factors, including age and AJCC, were significant in the test cohort. In the literature, multiple retrospective studies showed contradicting results regarding the prognostic value of pCR [57,58]. The previously mentioned meta-analysis encompassing 1663 predominantly retrospectively assessed patients demonstrated significant association of OS independent of the treatment modality [10]. Regarding the small patient numbers, we therefore see pCR as a valid endpoint for this exploratory work. Future larger studies should assess direct predictions of OS.

In many ways, delta-radiomics is a concept similar to assessing treatment response via changes in tumor volume by the RECIST criteria [50]. In contrast to the delta radiomic models, *RECIST*, *Delta-Volume,* and *Post-Volume* did not show similar predictive performances in the test set. Similar observations have been previously described [11,32]. *Pre-Volume*, however, appeared to be a stable predictor for pCR. The finding that radiomic features had better predictive performance than volume alone is of interest because radiomic features have been criticized as overly correlated to the VOI volume [59]. Delta-radiomic features may capture radiation-induced biologic changes occurring within the STS such as infarction, necrosis, fibrosis, or hyalinization [60].

In other cancer types, pCR plays an increasing role in treatment personalization. In breast cancer, certain pre-stratified patients without pCR after neoadjuvant Ctx receive additional adjuvant Ctx [8]. Future directions may harness imaging-based response assessment prior to surgery to identify those patients with STS who may benefit from additional neoadjuvant therapeutic escalation.

We also evaluated the transferability of the developed models to peri-CT delta-radiomic features. A high predictive performance could be observed. These results may be overly optimistic as all patients receiving CTx were in the training set. Still, this exploratory analysis suggests that the transfer of radiomic models to different cytotoxic therapies should be investigated.

Pathological response in clinical practice is evaluated on representative samples of the STS. This procedure constitutes a compromise between accurate response estimation and time expenditure [33]. As consequence, the determination of viable cells may be prone to a certain sampling error. By using a binarized endpoint, this risk may be reduced as minor numerical deviations far from the cut-point do not affect the result. Still, this uncertainty in the endpoint constitutes a limitation of this approach. Moreover, other pathological measures such as hyalinization were not available for

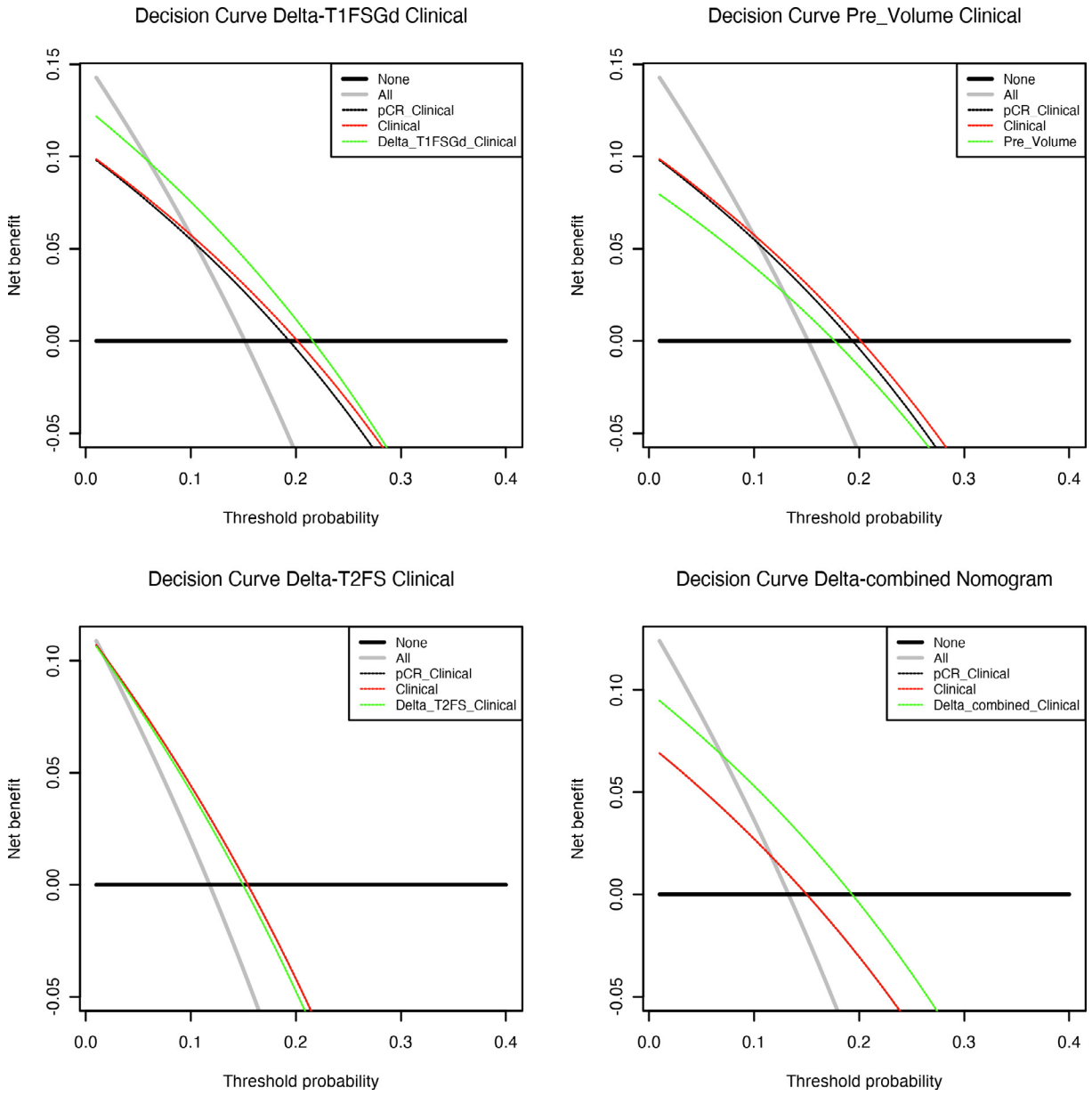

Fig. 4. Decision curve analysis of delta-radiomics multivariate models. Decision curve analysis was performed comparing the net benefit of the delta-radiomic multivariate models with the clinical model (AJCC+age) and the pCR multivariate model in the testing set. The net benefit is calculated by subtracting the proportion of false-positive patients from the proportion of true-positive patients, weighted by the relative harm of a false-positive and false-negative result [52]. The threshold probability was calculated for death after three years. The two extreme strategies "treat all" and "treat none" are displayed as a reference. A decision model shows a clinical benefit if the decision curve shows larger net benefit than both reference strategies. In the panel right panel in the second row "Clinical" and "pCR_Clinical" are overlapping. Abbreviations: AJCC: American Joint Committee on Cancer staging system 8th edition, AUC: area under the ROC curve, Delta: delta-radiomics models, pCR: pathological complete response, T1FSGd: contrast-enhanced T1-weighted fat saturated, T2FS: T2-weighted fat saturated.

all patients among both institutions, but constitute interesting alternative response markers [60].

This analysis bears several additional limitations. First, the retrospective nature of this study may be a reason for potential bias in data selection [61]-. Second, the heterogeneity in treatment regimens impaired therapy-specific modeling potentially reducing performance. Third, patient numbers were overall low, especially in the test set. Large standard deviations made direct comparisons between models difficult and impeded interpretability of multivariate models, especially in the context of the imbalanced outcome measure. For instance, in the test set none of the known prognostic factors was significantly associated with OS. Still, this study remains the largest reported delta-radiomic analysis in STS patients. Fourth, technical variations in image acquisition between cohorts and between the two timepoints used for delta-radiomic feature calculation may have hindered better performances and generalizability. Finally, the optimal approach to calculate deltaradiomic features is unknown. We decided for absolute subtraction of feature values as performed in previous studies, but the assessment of the relative change may be an alternative method. A prospective trial, with pre-defined acquisition protocols optimized for feature reproducibility, may enable better response evaluation.

To conclude, we could demonstrate prediction of pathological complete response using the delta-radiomics principle in STS

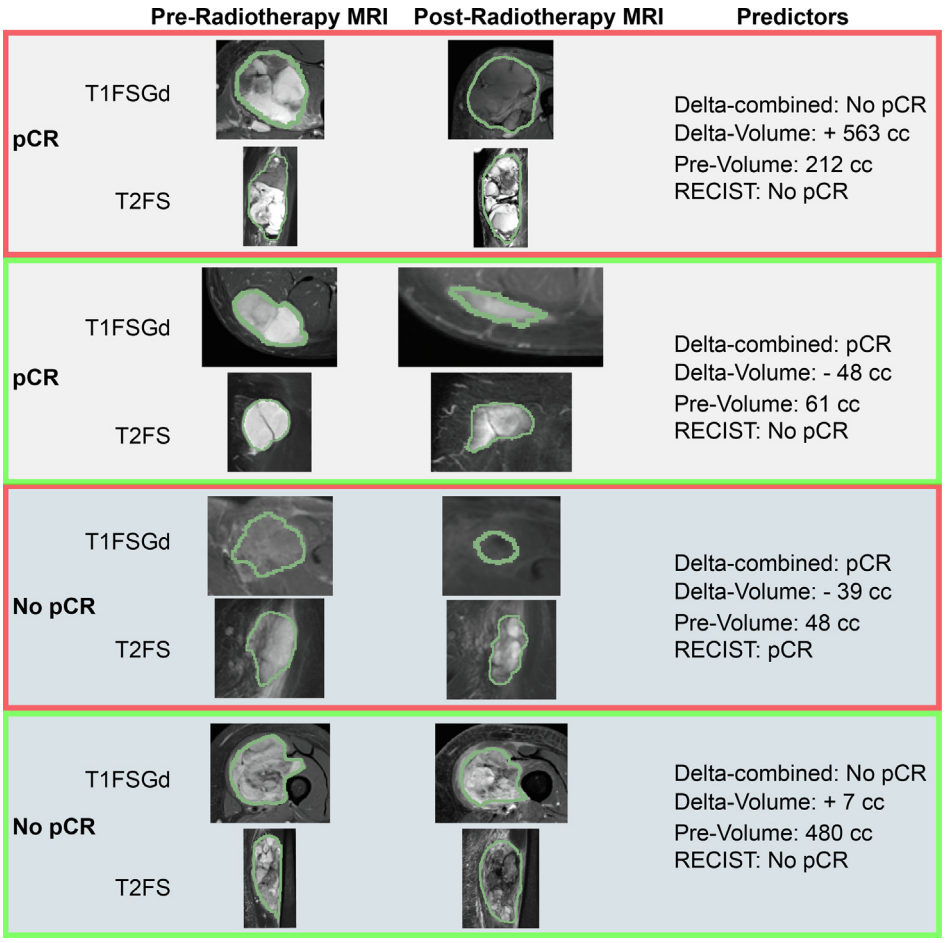

Fig. 5. Exemplary patient cases. Exemplary patient cases within the test cohort. Green and red frames represent correct or incorrect classification by the exemplary Delta-combined model. First patient: G3 undifferentiated pleomorphic sarcoma, second patient: G1 myxoid sarcoma, third patient: G2 monophasic synovialsarcoma, fourth patient: G3 myxofibrosarcoma. Abbreviations: Delta: delta-radiomics models, pCR: pathological complete response, T1FSGd: contrast-enhanced T1-weighted fat saturated, T2FS: T2-weighted fat saturated.

patients. Delta-radiomic models achieved better reproducibility than single timepoint radiomic models, RECIST, or the peri-therapeutic volume change and was associated with overall survival. The models also functioned in patients using MRIs obtained before and after chemotherapy. We conclude, that delta-radiomic features may capture radiation-induced biological changes and may function as a treatment response biomarker.

Funding

This work was funded by an ESTRO mobility grant and a physician-scientist program of the Helmholtz Zentrum München (JCP).

Declaration of Competing Interest

DH received research funding from GE Healthcare, Philips Healthcare, Canon Medical Systems USA, and Siemens Healthineers. PL reports, within the submitted work, grants/sponsored research agreements from radiomics SA, PL has minority shares in the company radiomics SA, he is co-inventor of two issued patents with royalties on radiomics (PCT/NL2014/050248, PCT/NL2014/050728) licensed to Radiomics SA, one non-patented invention (softwares) licensed to Radiomics SA and two non-issues, non-licensed patents on Deep Learning-Radiomics (n° 17123816, PCT/NL/2020/050794). PL confirms that none of the above entities or funding was involved in the preparation of this paper. HW has minority shares in the company radiomics SA. The other authors declare no potential conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2021.08.023.

References

[1] Clark MA, Fisher C, Judson I, Thomas JM. Soft-tissue sarcomas in adults. N Engl J Med 2005;353:701–11. https://doi.org/10.1056/NEJMra041866.
[2] Gutierrez JC, Perez EA, Franceschi D, Moffat FL, Livingstone AS, Koniaris LG. Outcomes for soft-tissue sarcoma in 8249 cases from a large state cancer registry. J Surg Res 2007;141:105–14. https://doi.org/10.1016/j.jss.2007.02.026.
[3] O'Sullivan B, Davis AM, Turcotte R, Bell R, Catton C, Chabot P, et al. Preoperative verses postoperative radiotherapy in soft-tissue sarcoma of the limbs: a randomised trial. Lancet 2002;359:2235–41. https://doi.org/10.1016/S0140-6736(02)09292-9.
[4] Peeken JC, Knie C, Kessel KA, Habermehl D, Kampfer S, Dapper H, et al. Neoadjuvant image-guided helical intensity modulated radiotherapy of extremity sarcomas – a single center experience. Radiat Oncol 2019;14. https://doi.org/10.1186/s13014-019-1207-2.
[5] Lartigau E, Kantor G, Taieb S, Vilain MO, Ceugnart L, Lagarde P, et al. Definitions of target volumes in soft tissue sarcomas of the extremities. Cancer Radiother 2001;5:695–703.
[6] Woll PJ, Reichardt P, Le Cesne A, Bonvalot S, Azzarelli A, Hoekstra HJ, et al. Adjuvant chemotherapy with doxorubicin, ifosfamide, and lenograstim for resected soft-tissue sarcoma (EORTC 62931): A multicentre randomised

controlled trial. Lancet Oncol 2012;13:1045–54. https://doi.org/10.1016/S1470-2045(12)70346-7.
[7] Gronchi A, Ferrari S, Quagliuolo V, Broto JM, Pousa AL, Grignani G, et al. Histotype-tailored neoadjuvant chemotherapy versus standard chemotherapy in patients with high-risk soft-tissue sarcomas (ISG-STS 1001): an international, open-label, randomised, controlled, phase 3, multicentre trial. Lancet Oncol 2017;18:812–22. https://doi.org/10.1016/S1470-2045(17)30334-0.
[8] Masuda N, Lee S-J, Ohtani S, Im Y-H, Lee E-S, Yokota I, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. N Engl J Med 2017;376:2147–59. https://doi.org/10.1056/NEJMoa1612645.
[9] Bonvalot S, Rutkowski PL, Thariat J, Carrère S, Ducassou A, Sunyach M, et al. NBTXR3, a first-in-class radioenhancer hafnium oxide nanoparticle, plus radiotherapy versus radiotherapy alone in patients with locally advanced soft-tissue sarcoma (Act In Sarc): a multicentre, phase 2–3, randomised , controlled trial. 2019;20. doi:10.1016/S1470-2045(19)30326-2.
[10] Salah S, Lewin J, Amir E, Abdul RA. Tumor necrosis and clinical outcomes following neoadjuvant therapy in soft tissue sarcoma: A systematic review and meta-analysis. Cancer Treat Rev 2018;69:1–10. https://doi.org/10.1016/j.ctrv.2018.05.007.
[11] Gennaro N, Reijers S, Bruining A, Messiou C, Haas R, Colombo P, et al. Imaging response evaluation after neoadjuvant treatment in soft tissue sarcomas: Where do we stand? Crit Rev Oncol Hematol 2021;160:103309. https://doi.org/10.1016/j.critrevonc.2021.103309.
[12] Nussbaum DP. Nanoparticle augmentation of radiotherapy in sarcoma. Lancet Oncol 2019;20:1046–8. https://doi.org/10.1016/S1470-2045(19)30392-4.
[13] Peeken JC, Nüsslin F, Combs SE. "Radio-oncomics" - The potential of radiomics in radiation oncology„Radio-oncomics": Das Potenzial von Radiomics in der Strahlenonkologie. Strahlentherapie Und Onkol 2017;193:767–79. https://doi.org/10.1007/s00066-017-1175-0.
[14] Lambin P, van Stiphout RGPM, Starmans MHW, Rios-Velazquez E, Nalbantov G, Aerts HJWL, et al. Predicting outcomes in radiation oncology–multifactorial decision support systems. Nat Rev Clin Oncol 2013;10:27–40. https://doi.org/10.1038/nrclinonc.2012.196.
[15] Peeken JC, Wiestler B, Combs SE. The potential of radiomics in clinical application. In: Debus J, Schober O, Kiessling F, editors. Image Guide. Radiooncology, Springer Berlin Heidelberg; 2020.
[16] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5. https://doi.org/10.1038/ncomms5006.
[17] Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. Cancer Res 2017;77:3922–30. https://doi.org/10.1158/0008-5472.CAN-17-0122.
[18] Diehn M, Nardini C, Wang DS, McGovern S, Jayaraman M, Liang Y, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. Proc Natl Acad Sci U S A 2008;105:5213–8. https://doi.org/10.1073/pnas.0801279105.
[19] Peeken JC, Shouman MA, Kroenke M, Rauscher I, Maurer T, Gschwend JE, et al. A CT-based radiomics model to detect prostate cancer lymph node metastases in PSMA radioguided surgery patients. Eur J Nucl Med Mol Imaging 2020;47:2968–77. https://doi.org/10.1007/s00259-020-04864-1.
[20] Peeken JC, Molina-Romero M, Diehl C, Menze BH, Straube C, Meyer B, et al. Deep learning derived tumor infiltration maps for personalized target definition in Glioblastoma radiotherapy. Radiother Oncol 2019;138:166–72. https://doi.org/10.1016/j.radonc.2019.06.031.
[21] Leger S, Zwanenburg A, Leger K, Lohaus F, Linge A, Schreiber A, et al. Comprehensive analysis of tumour sub-volumes for radiomic risk modelling in locally advanced HNSCC. Cancers (Basel) 2020;12:3047. https://doi.org/10.3390/cancers12103047.
[22] Navarro F, Dapper H, Asadpour R, Knebel C, Spraker MB, Schwarze V, et al. Development and external validation of deep-learning-based tumor grading models in soft-tissue sarcoma patients using MR imaging. Cancers (Basel) 2021;13:2866. https://doi.org/10.3390/cancers13122866.
[23] Peeken JC, Bernhofer M, Spraker MB, Pfeiffer D, Devecka M, Thamer A, et al. CT-based radiomic features predict tumor grading and have prognostic value in patients with soft tissue sarcomas treated with neoadjuvant radiation therapy. Radiother Oncol 2019;135:187–96. https://doi.org/10.1016/j.radonc.2019.01.004.
[24] Spraker MB, Wootton LS, Hippe DS, Ball KC, Peeken JC, Macomber MW, et al. MRI radiomic features are independently associated with overall survival in soft tissue sarcoma. Adv Radiat Oncol 2019;4:413–21. https://doi.org/10.1016/j.adro.2019.02.003.
[25] Navarro F, Shit S, Ezhov I, Paetzold J, Gafita A, Peeken JC, et al. Shape-aware complementary-task learning for multi-organ segmentation. MLMI Work. 2019 MICCAI, 2019, p. 620–7. doi:10.1007/978-3-030-32692-0_71.
[26] Vallières M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 2015;60:5471–96. https://doi.org/10.1088/0031-9155/60/14/5471.
[27] Crombé A, Le Loarer F, Sitbon M, Italiano A, Stoeckle E, Buy X, et al. Can radiomics improve the prediction of metastatic relapse of myxoid/round cell liposarcomas? Eur Radiol 2020;30:2413–24. https://doi.org/10.1007/s00330-019-06562-5.
[28] Peeken JC, Neumann J, Asadpour R, Leonhardt Y, Moreira JR, Hippe DS, et al. Prognostic assessment in high-grade soft-tissue sarcoma patients: A comparison of semantic image analysis and radiomics. Cancers (Basel) 2021;13:1–17. https://doi.org/10.3390/cancers13081929.
[29] Specht L, Yahalom J, Illidge T, Berthelsen AK, Constine LS, Eich HT, et al. Modern radiation therapy for Hodgkin lymphoma: Field and dose guidelines from the international lymphoma radiation oncology group (ILROG). Int J Radiat Oncol Biol Phys 2014;89:854–62. https://doi.org/10.1016/j.ijrobp.2013.05.005.
[30] Fave X, Zhang L, Yang J, Mackin D, Balter P, Gomez D, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. Sci Rep 2017;7. https://doi.org/10.1038/s41598-017-00665-z.
[31] van Timmeren JE, Leijenaar RTH, van Elmpt W, Reymen B, Lambin P. Feature selection methodology for longitudinal cone-beam CT radiomics. Acta Oncol (Madr) 2017;56:1537–43. https://doi.org/10.1080/0284186X.2017.1350285.
[32] Crombé A, Périer C, Kind M, De Senneville BD, Le Loarer F, Italiano A, et al. T2-based MRI Delta-radiomics improve response prediction in soft-tissue sarcomas treated by neoadjuvant chemotherapy. J Magn Reson Imaging 2019;50:497–510. https://doi.org/10.1002/jmri.v50.210.1002/jmri.26589.
[33] Wardelmann E, Haas RL, Bovée JVMG, Terrier Ph, Lazar A, Messiou C, et al. Evaluation of response after neoadjuvant treatment in soft tissue sarcomas; the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC-STBSG) recommendations for pathological examination and reporting. Eur J Cancer 2016;53:84–95. https://doi.org/10.1016/j.ejca.2015.09.021.
[34] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Ann Intern Med 2015;162:W1–W73. https://doi.org/10.7326/M14-0698.
[35] Fedorov A, Beichel R, Kalphaty-Cramer J, Finet J, Fillion-Robbin J-C, Pujol S, et al. 3D slicers as an image computing platform for thw quantitative imaging network. Magn Reson Imaging 2012;30:1323–41. https://doi.org/10.1016/j.mri.2012.05.001.3D.
[36] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339.
[37] Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020:191145. doi:10.1148/radiol.2020191145.
[38] Lucia F, Visvikis D, Vallières M, Desseroit M-C, Miranda O, Robin P, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019;46:864–77. https://doi.org/10.1007/s00259-018-4231-9.
[39] Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of a method to compensate multicenter effects affecting CT radiomics. Radiology 2019;291:53–9. https://doi.org/10.1148/radiol.2019182023.
[40] Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. NeuroImage Harmonization of multi-site diffusion tensor imaging data. Neuroimage 2017;161:149–70. https://doi.org/10.1016/j.neuroimage.2017.08.047.
[41] Steiger P, Sood R. How can radiomics be consistently applied across imagers and institutions? Radiology 2019;291:60–1. https://doi.org/10.1148/radiol.2019190051.
[42] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118–27. https://doi.org/10.1093/biostatistics/kxj037.
[43] Kursa MB, Rudnicki WR. Feature selection with the boruta package. J Stat Softw 2010;36:1–13. https://doi.org/10.18637/jss.v036.i11.
[44] Wu G, Woodruff HC, Shen J, Refaee T, Sanduleanu S, Abdalla I, et al. Diagnosis of invasive lung adenocarcinoma based on chest CT radiomic features of part-solid pulmonary nodules: a multicenter study. Radiology 2020:192431. doi:10.1148/radiol.2020192431.
[45] Deist TM, Dankers FJWM, Valdes G, Wijsman R, Hsu I-C, Oberije C, et al. Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers. Med Phys 2018;45:3449–59. https://doi.org/10.1002/mp.2018.45.issue-710.1002/mp.12967.
[46] Peeken JC, Bernhofer M, Wiestler B, Goldberg T, Cremers D, Rost B, et al. Radiomics in radiooncology - Challenging the medical physicist. Phys Medica 2018;48:27–36. https://doi.org/10.1016/j.ejmp.2018.03.012.
[47] Waldron L, Pintilie M, Tsao MS, Shepherd FA, Huttenhower C, Jurisica I. Optimized application of penalized regression methods to diverse genomic data. Bioinformatics 2011;27:3399–406. https://doi.org/10.1093/bioinformatics/btr591.
[48] Friedman JH. Greedy function approximation: a gradient boosting machine. IMS Reitz Lect 1999:1–39.
[49] Breiman L. Random forests. Mach Learn 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
[50] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–47. https://doi.org/10.1016/j.ejca.2008.10.026.
[51] Peeken JC, Spraker MB, Knebel C, Dapper H, Pfeiffer D, Devecka M, et al. Tumor grading of soft tissue sarcomas using MRI-based radiomics. EBioMedicine 2019;48:332–40. https://doi.org/10.1016/j.ebiom.2019.08.059.

[52] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak 2008;26:565–74. https://doi.org/10.1177/0272989X06295361.Decisio.

[53] Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: From utopia to empirical data. J Clin Epidemiol 2016;74:167–76. https://doi.org/10.1016/j.jclinepi.2015.12.005.

[54] Gao Y, Kalbasi A, Hsu W, Ruan D, Fu J, Shao J, et al. Treatment effect prediction for sarcoma patients treated with preoperative radiotherapy using radiomics features from longitudinal diffusion-weighted MRIs. Phys Med Biol 2020;65. doi:10.1088/1361-6560/ab9e58.

[55] Crombé A, Fadli D, Buy X, Italiano A, Saut O, Kind M. High-grade soft-tissue sarcomas: can optimizing dynamic contrast-enhanced MRI postprocessing improve prognostic radiomics models? J Magn Reson Imaging 2020;52:282–97. https://doi.org/10.1002/jmri.v52.110.1002/jmri.27040.

[56] Lin P, Yang PF, Chen S, Shao YY, Xu L, Wu Y, et al. A Delta-radiomics model for preoperative evaluation of Neoadjuvant chemotherapy response in high-grade osteosarcoma. Cancer Imaging 2020;20. https://doi.org/10.1186/s40644-019-0283-8.

[57] Eilber FC, Rosen G, Eckardt J, Forscher C, Nelson SD, Selch M, et al. Treatment-induced pathologic necrosis: A predictor of local recurrence and survival in patients receiving neoadjuvant therapy for high-grade extremity soft tissue sarcomas. J Clin Oncol 2001;19:3203–9. https://doi.org/10.1200/JCO.2001.19.13.3203.

[58] Mullen JT, Hornicek FJ, Harmon DC, Raskin KA, Chen YL, Szymonifka J, et al. Prognostic significance of treatment-induced pathologic necrosis in extremity and truncal soft tissue sarcoma after neoadjuvant chemoradiotherapy. Cancer 2014;120:3676–82. https://doi.org/10.1002/cncr.28945.

[59] Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: The need for safeguards. Radiother Oncol 2019;130:2–9. https://doi.org/10.1016/j.radonc.2018.10.027.

[60] Schaefer I-M, Hornick JL, Barysauskas CM, Raut CP, Patel SA, Royce TJ, et al. Histologic appearance after preoperative radiation therapy for soft tissue sarcoma: assessment of the European Organization for Research and Treatment of CancereSoft Tissue and Bone Sarcoma Group Response Score. Int J Radiat Oncol Biol Phys 2017;98:375–83. https://doi.org/10.1016/j.ijrobp.2017.02.087.

[61] Sica GT. Bias in research studies. Radiology 2006;238:780–9. https://doi.org/10.1148/radiol.2383041109.