Supplementary

Appendix 1 Radiomics quality score results of this study

Image protocol quality - well-documented image protocols (for example, contrast, slice thickness, energy, etc.) and/or usage of public image protocols allow reproducibility/replicability
- ☑ protocols well documented
- ☐ public protocol used
- ☐ none

Multiple segmentations - possible actions are: segmentation by different physicians/algorithms/software, perturbing segmentations by (random) noise, segmentation at different breathing cycles. Analyse feature robustness to segmentation variabilities
- ◉ yes
- ○ no

Phantom study on all scanners - detect inter-scanner differences and vendor-dependent features. Analyse feature robustness to these sources of variability
- ○ yes
- ◉ no

Imaging at multiple time points - collect images of individuals at additional time points. Analyse feature robustness to temporal variabilities (for example, organ movement, organ expansion/shrinkage)
- ○ yes
- ◉ no

Feature reduction or adjustment for multiple testing - decreases the risk of overfitting. Overfitting is inevitable if the number of features exceeds the number of samples. Consider feature robustness when selecting features
- ◉ Either measure is implemented
- ○ Neither measure is implemented

Multivariable analysis with non radiomics features (for example, EGFR mutation) - is expected to provide a more holistic model. Permits correlating/inferencing between radiomics and non radiomics features
- ◉ yes
- ○ no

Detect and discuss biological correlates - demonstration of phenotypic differences (possibly associated with underlying gene–protein expression patterns) deepens understanding of radiomics and biology
- ○ yes
- ◉ no

Cut-off analyses - determine risk groups by either the median, a previously published cut-off or report a continuous risk variable. Reduces the risk of reporting overly optimistic results
- ◉ yes
- ○ no

Discrimination statistics - report discrimination statistics (for example, C-statistic, ROC curve, AUC) and their statistical significance (for example, p-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)
- ☑ a discrimination statistic and its statistical significance are reported
- ☑ a resampling method technique is also applied
- ☐ none

Calibration statistics - report calibration statistics (for example, Calibration-in-the-large/slope, calibration plots) and their statistical significance (for example, P-values, confidence intervals). One can also apply resampling method (for example, bootstrapping, cross-validation)
- ☑ a calibration statistic and its statistical significance are reported
- ☑ a resampling method technique is applied
- ☐ none

Prospective study registered in a trial database - provides the highest level of evidence supporting the clinical validity and usefulness of the radiomics biomarker
- ○ yes
- ◉ no

Validation - the validation is performed without retraining and without adaptation of the cut-off value, provides crucial information with regard to credible clinical performance
- ☐ No validation
- ☐ validation is based on a dataset from the same institute
- ☐ validation is based on a dataset from another institute
- ☑ validation is based on two datasets from two distinct institutes
- ☐ the study validates a previously published signature
- ☐ validation is based on three or more datasets from distinct institutes

Comparison to 'gold standard' - assess the extent to which the model agrees with/is superior to the current 'gold standard' method (for example, TNM-staging for survival prediction). This comparison shows the added value of radiomics
- ○ yes
- ◉ no

Potential clinical utility - report on the current and potential application of the model in a clinical setting (for example, decision curve analysis).
- ◉ yes
- ○ no

Cost-effectiveness analysis - report on the cost-effectiveness of the clinical application (for example, QALYs generated)
- ○ yes
- ◉ no

Open science and data - make code and data publicly available. Open science facilitates knowledge transfer and reproducibility of the study
- ☐ scans are open source
- ☐ region of interest segmentations are open source
- ☑ the code is open sourced
- ☑ radiomics features are calculated on a set of representative ROIs and the calculated features and representative ROIs are open source

Total score

19
(52.78%)

https://dx.doi.org/10.21037/qims-24-35

Appendix 2 Computational formulas for RS-CMP, RS-NP, RS-NCP, and RS-Triphasic

1. **RS-CMP=**–1.3951 – 0.0762*lbp.3D.m1_firstorder_Skewness_A + 0.1142*log.sigma.1.0.mm.3D_glszm_SmallAreaEmphasis_A – 0.0083*log.sigma.2.0.mm.3D_glcm_Imc2_A + 0.0684*log.sigma.4.0.mm.3D_glszm_LargeAreaLowGrayLevelEmphasis_A + 0.1061*log.sigma.5.0.mm.3D_firstorder_Maximum_A + 0.0025*log.sigma.5.0.mm.3D_glcm_Idmn_A + 0.2346*log.sigma.5.0.mm.3D_glrlm_RunVariance_A –0.1811*logarithm_glcm_InverseVariance_A + 0.0286*logarithm_ngtdm_Complexity_A – 0.0791*original_shape_Flatness_A – 0.1274*wavelet.HH_firstorder_Median_A + 0.0221*wavelet.LH_glcm_Idn_A –0.4767*wavelet.LH_gldm_SmallDependenceLowGrayLevelEmphasis_A + 0.0072*wavelet.LL_glcm_JointEnergy_A + 0.0667*wavelet.LL_glcm_MaximumProbability_A + 0.0275*wavelet.LL_glrlm_RunVariance_A

2. RS-NP=–1.3208 + 0.1426*exponential_glszm_LargeAreaEmphasis_V + 4.55E–05*exponential_glszm_LargeAreaLowGrayLevelEmphasis_V + 0.0322*lbp.3D.k_glcm_ClusterProminence_V + 0.0147*lbp.3D.k_glszm_LargeAreaHighGrayLevelEmphasis_V + 0.0306*log.sigma.1.0.mm.3D_glszm_LargeAreaLowGrayLevelEmphasis_V + 0.105*log.sigma.2.0.mm.3D_firstorder_Median_V + 0.0157*log.sigma.4.0.mm.3D_glszm_LargeAreaLowGrayLevelEmphasis_V + 0.1125*log.sigma.5.0.mm.3D_firstorder_Maximum_V + 0.0949*logarithm_ngtdm_Busyness_V + 0.1612*square_glcm_Idmn_V + 0.1136*wavelet.LH_gldm_DependenceNonUniformityNormalized_V – 0.0708*wavelet.LH_gldm_DependenceVariance_V –0.0466*wavelet.LL_firstorder_Minimum_V + 0.0558*wavelet.LL_glszm_LargeAreaEmphasis_V

3. RS-NCP=–1.3276 + 0.0399*lbp.2D_glszm_ZoneVariance_P + 4.58E–16*lbp.3D.m1_glszm_ZoneVariance_P + 3.36E–16*lbp.3D.m2_glszm_ZoneVariance_P + 0.0675*log.sigma.1.0.mm.3D_firstorder_Kurtosis_P + 0.1227*log.sigma.3.0.mm.3D_glcm_Idmn_P + 0.0686*logarithm_glcm_ClusterProminence_P + 0.0195*original_glcm_Idmn_P + 0.1169*original_glcm_Idn_P – 0.2946*original_shape_Sphericity_P + 0.0917*wavelet.HH_glszm_ZoneVariance_P + 0.0467*wavelet.LH_firstorder_Mean_P

4. RS-Triphasic = –1.4321 + 0.0467*lbp.3D.k_ngtdm_Contrast_A – 0.0095*lbp.3D.m1_firstorder_Skewness_A –0.0265*lbp.3D.m2_firstorder_10Percentile_A + 0.1603*log.sigma.1.0.mm.3D_glszm_SmallAreaEmphasis_A + 0.0299*log.sigma.4.0.mm.3D_glszm_LargeAreaLowGrayLevelEmphasis_A + 0.2068*log.sigma.5.0.mm.3D_glrlm_RunVariance_A –0.2185*logarithm_glcm_InverseVariance_A –0.0539*wavelet.HH_firstorder_Median_A –0.3168*wavelet.LH_gldm_SmallDependenceLowGrayLevelEmphasis_A + 0.017*lbp.3D.k_glcm_ClusterProminence_V + 0.1833*wavelet.LH_gldm_DependenceNonUniformityNormalized_V –0.0331*wavelet.LL_firstorder_Minimum_V + 0.0043*wavelet.LL_glrlm_ShortRunHighGrayLevelEmphasis_V + 0.0637*gradient_glszm_SmallAreaLowGrayLevelEmphasis_P + 0.0693*log.sigma.1.0.mm.3D_firstorder_Kurtosis_P –0.0215*log.sigma.2.0.mm.3D_glszm_GrayLevelNonUniformityNormalized_P + 0.0066*log.sigma.5.0.mm.3D_firstorder_Entropy_P + 0.058*logarithm_glcm_ClusterProminence_P + 0.1107*logarithm_glcm_Correlation_P + 0.0837*original_glcm_MaximumProbability_P + 0.0973*original_gldm_DependenceVariance_P –0.261*original_shape_Sphericity_P – 0.0702*squareroot_glcm_ClusterShade_P + 0.0149*squareroot_glcm_Idn_P + 0.0446*wavelet.LH_firstorder_Mean_P

https://dx.doi.org/10.21037/qims-24-35

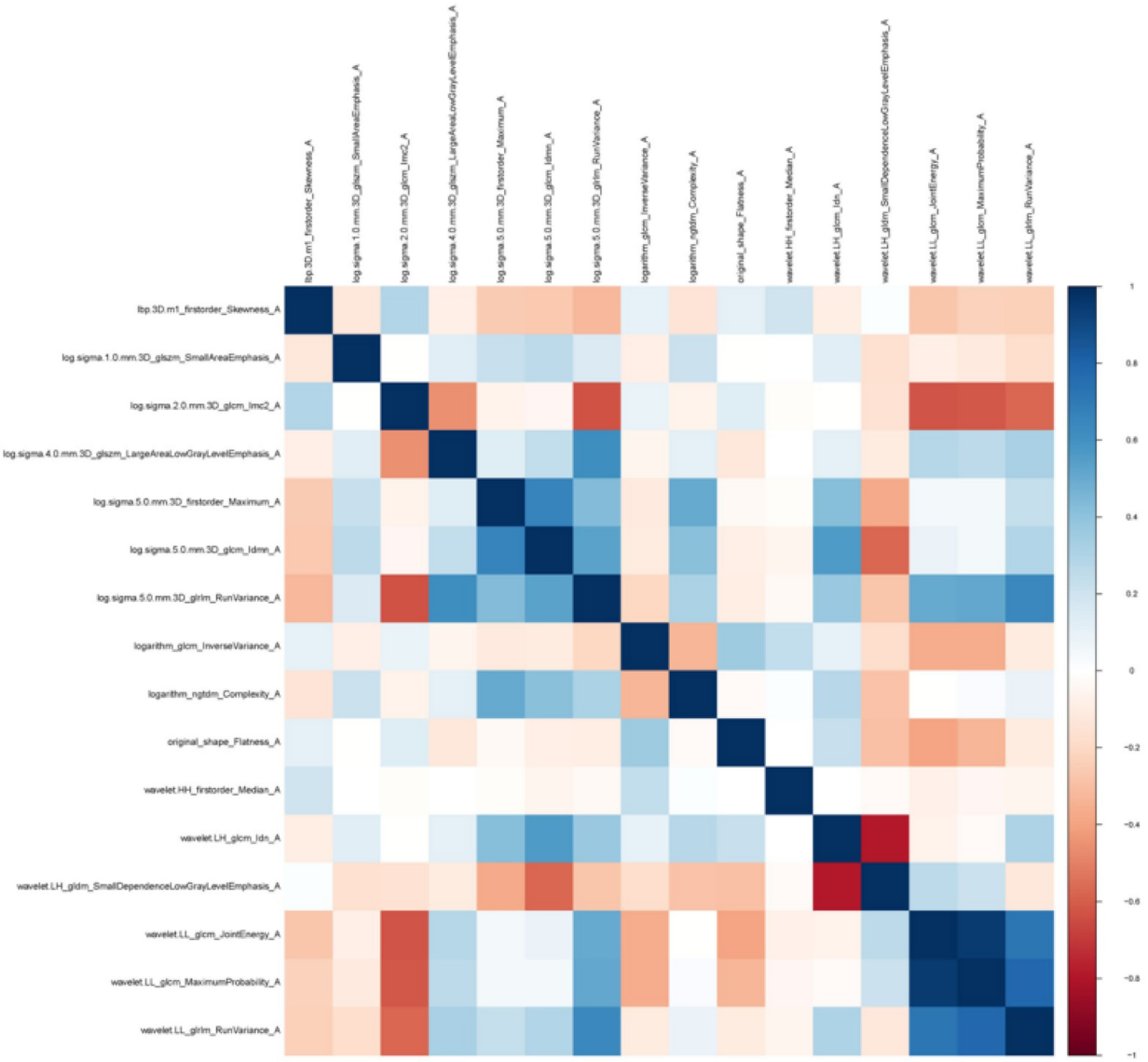

Figure S1 Correlation heatmap of the selected radiomics features from the CMP images. CMP, corticomedullary phase.

https://dx.doi.org/10.21037/qims-24-35

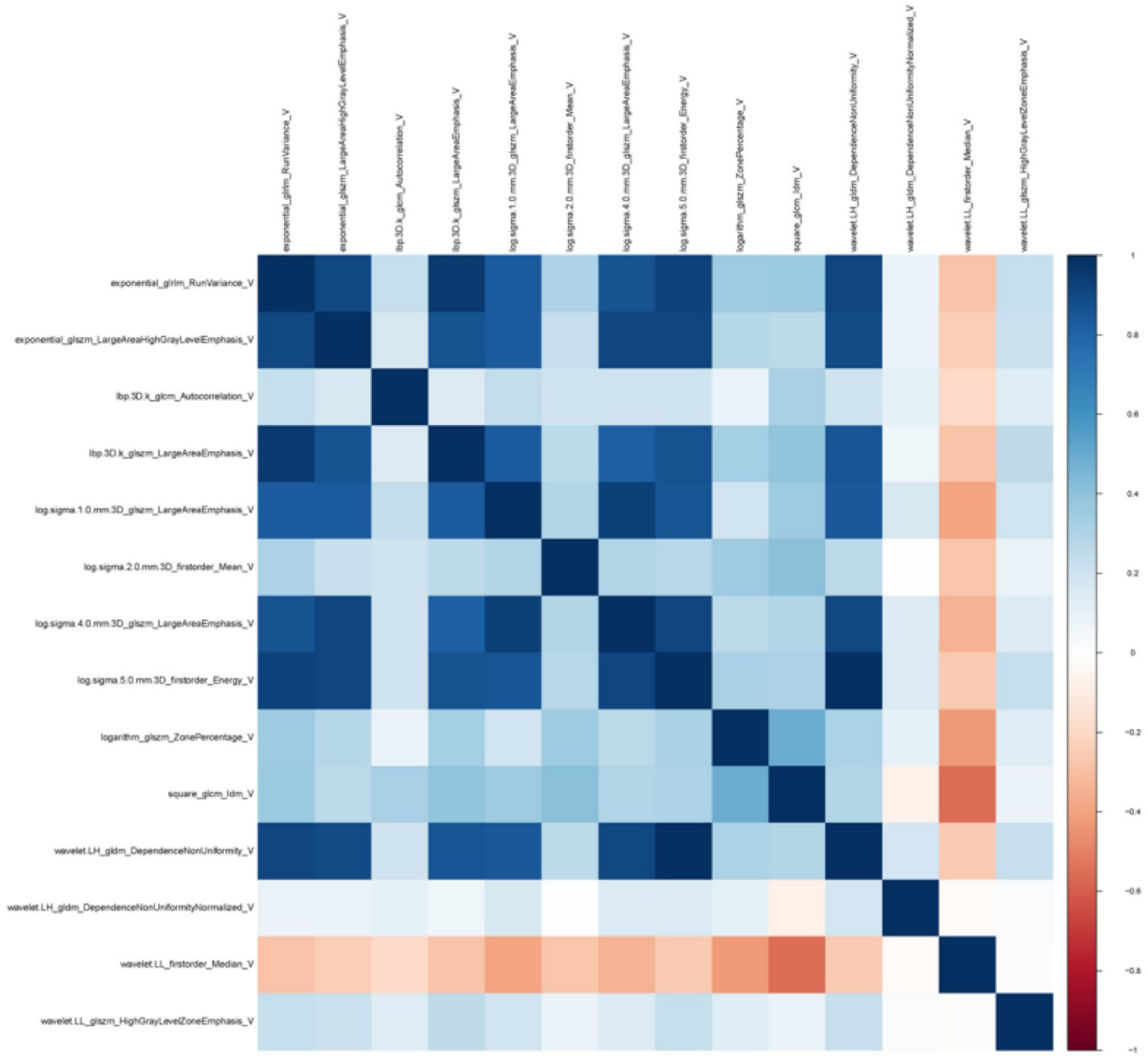

Figure S2 Correlation heatmap of the selected radiomics features from the NP images. NP, nephrographic phase.

https://dx.doi.org/10.21037/qims-24-35

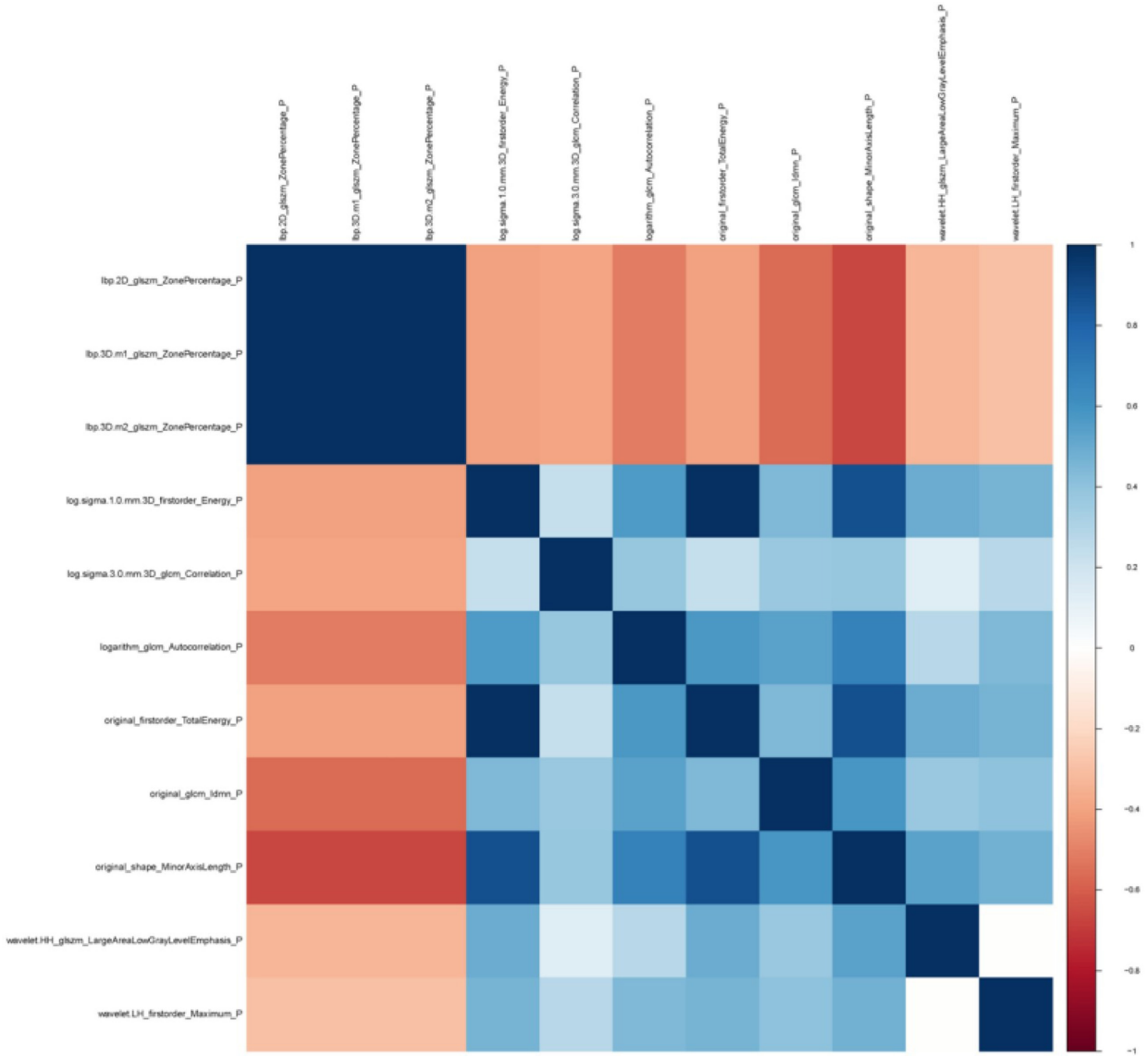

Figure S3 Correlation heatmap of the selected radiomics features from the NCP images. NCP, non-contrast phase.

https://dx.doi.org/10.21037/qims-24-35

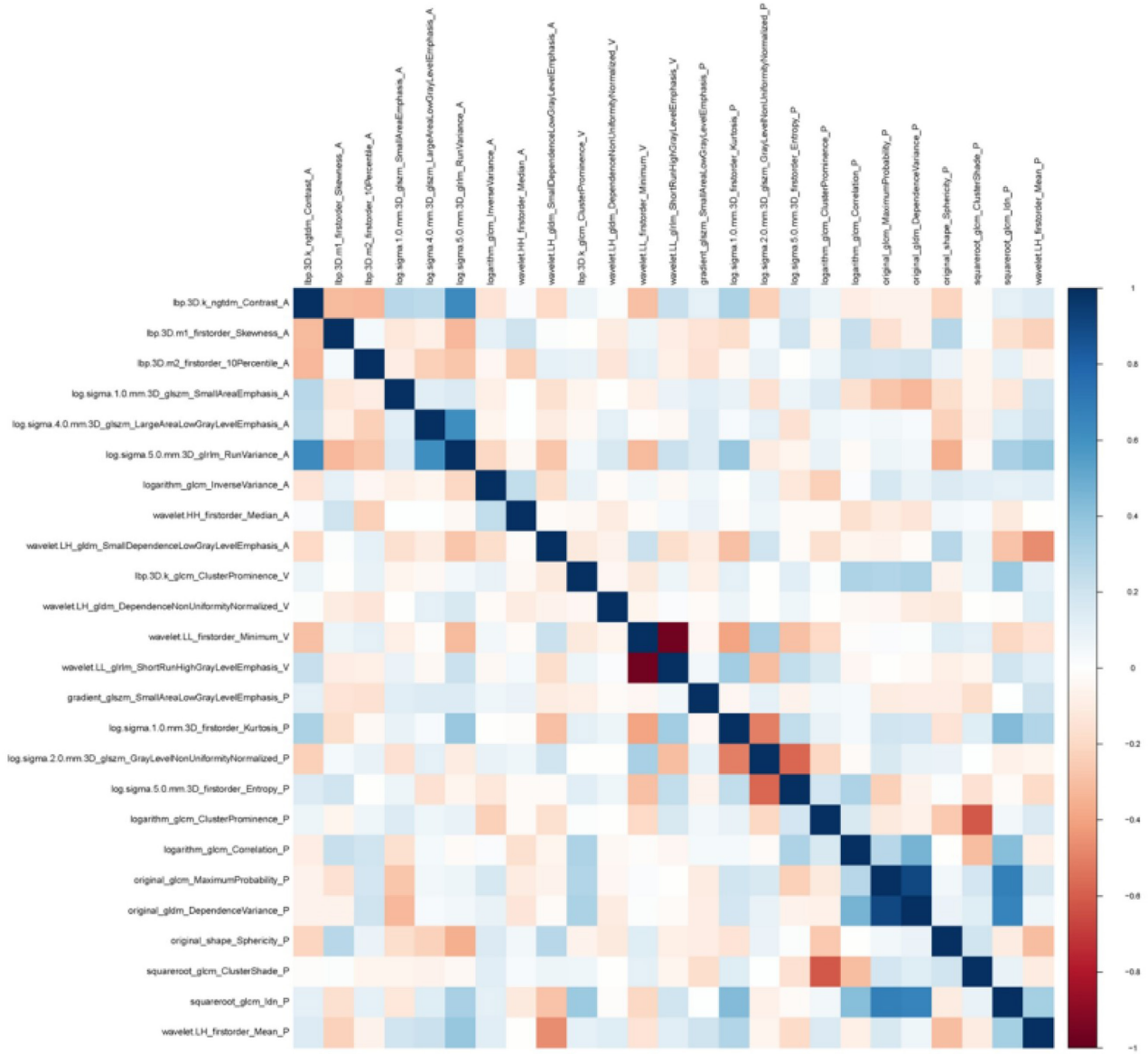

Figure S4 Correlation heatmap of the selected radiomics features the from triphasic images.

https://dx.doi.org/10.21037/qims-24-35

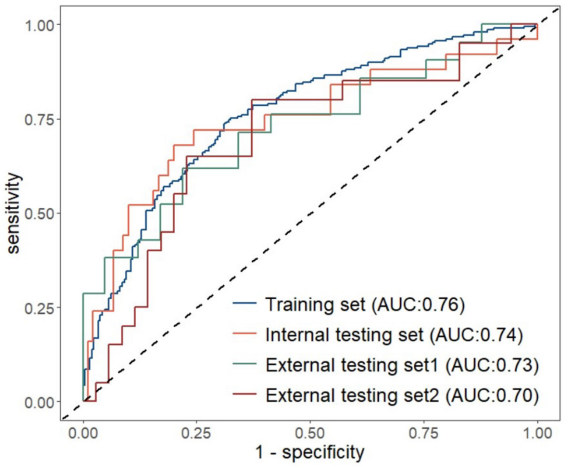

Figure S5 ROC curves of the clinical model in the training, and internal and external testing sets. ROC, receiver operating characteristic; AUC, area under the curve.

Table S1 CT acquisition parameters

Device	Discovery 750 HD, GE Healthcare	Siemens Sensation 16	SOMATOM Definition Flash system, Siemens Medical Systems
Sequence	Axial	Axial	Axial
Tube voltage (kVp)	120–140	120	110–120
Tube current (mAs)	220–300	150	50–150
Gantry rotation time (s)	0.5	0.5	0.5
Detector collimation (mm)	64×0.625	16×1.5	128×0.6
Matrix	512×512	512×512	512×512
Slice thickness (mm)	5.0	1.0	1.0
Hospitals	a, b, c	a, c	a, b

a, The First Affiliated Hospital of Chongqing Medical University; b, The Second Affiliated Hospital of Chongqing Medical University; c, Yongchuan Hospital of Chongqing Medical University. CT, computed tomography.

https://dx.doi.org/10.21037/qims-24-35

Table S2 Types of extracted radiomics features

First order statistics
Shape-based (3D)
Shape-based (2D)
Gray-level co-occurrence matrix (GLCM)
Gray-level run-length matrix (GLRLM)
Gray-level size-zone matrix (GLSZM)
Gray-level dependence matrix (GLDM)
Neighboring gray-tone difference matrix (NGTDM)

The radiomics features were extracted from both original images and derived images that were obtained by applying filters.

Table S3 Multivariable logistic regression results for the clinical variates

Variate	Full multivariate model			Reduced multivariate model		
	Coefficient	OR (95% CI)	P value	Coefficient	OR (95% CI)	P value
Intercept	0.28	1.32 (0.23–7.57)	0.76	–0.20	0.82 (0.27–2.50)	0.73
Age	–0.02	0.98 (0.96–1.00)	0.06	–0.02	0.98 (0.97–1.00)	0.06
Sex	–0.37	0.69 (0.46–1.06)	0.09	–0.38	0.68 (0.45–1.03)	0.07
BMI	–0.02	0.98 (0.92–1.04)	0.48	–	–	–
Intratumoral necrosis	–0.63	0.53 (0.33–0.87)	0.01	–0.62	0.54 (0.33–0.88)	0.01
Tumor size	0.37	1.45 (1.26–1.68)	<0.001	0.34	1.40 (1.26–1.55)	<0.001
Clinical T stage ≥2	–0.30	0.74 (0.34–1.63)	0.45	–	–	–
Clinical N1 stage	–1.29	0.28 (0.04–1.71)	0.17	–	–	–
Clinical M1 stage	15.41	–	0.98	–	–	–

OR, odds ratio; CI, confidence interval; BMI, body mass index.

https://dx.doi.org/10.21037/qims-24-35