

## Appendix 1

### *SDCT protocol and acquisition*

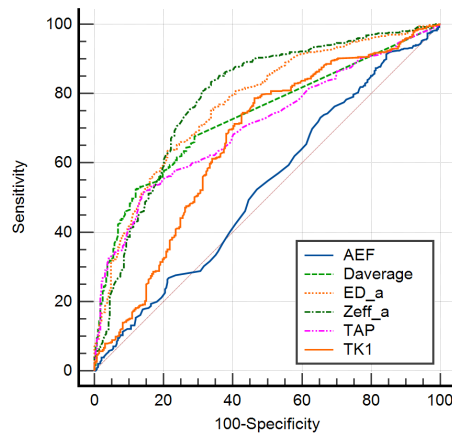
The scanning and reconstruction settings were as follows: a tube voltage of 120 kV, with automatic modulation of tube current, a noise index of 0.65, a rotation speed of 0.33 seconds per rotation, a pitch of 0.671, and a collimator width of 64 mm × 0.625 mm. The matrix size was 512×512. Spectral base images were reconstructed with a slice thickness of 1 mm and a 1 mm increment using iDose-level 3, utilizing two reconstruction kernels: standard (B) for mediastinal windows and Y-Detail (YB) for lung windows.

### *SDCT image analysis*

The following parameters were calculated: (I) non-contrast-enhanced CT values [in Hounsfield unit (HU)] at  $CT_{40\text{ keV}}$  and  $CT_{100\text{ keV}}$ ; (II) spectral curve slopes (slope) =  $|CT_{40\text{ keV}} - CT_{100\text{ keV}}| / (100 - 40)$ ; (III) NID = ID/ID aorta (where the aorta is either the thoracic aorta or subclavian artery in the same slice), to account for variations in patient hemodynamics and contrast agent distribution; and (IV) arterial enhancement fraction (AEF) =  $T1/T2 \times 100\%$ , where T1 and T2 represent the IDs in the arterial and venous phases, respectively.

## Appendix 2

ROC curves of different SDCT parameters and clinical indicators.

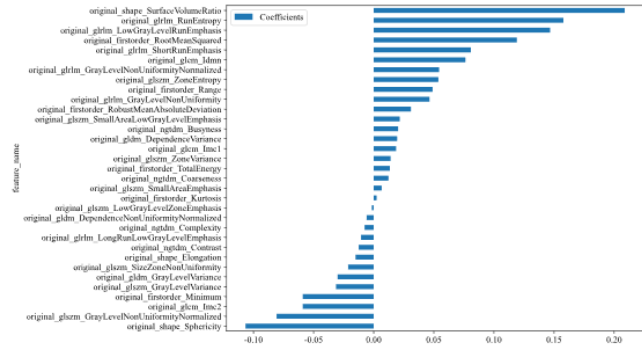


ROC, receiver operating characteristic; SDCT, spectral dual-layer detector computed tomography; AEF, arterial enhancement fraction; ED, electron density; a, arterial phase; Zeff, effective atomic number; TAP, tumor abnormal protein; TK1, thymidine kinase 1.

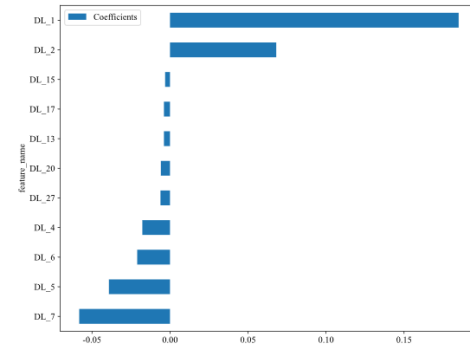
### Appendix 3

Histogram based on the radiomics score of the selected feature for the coefficient values of the non-zero features ultimately selected.

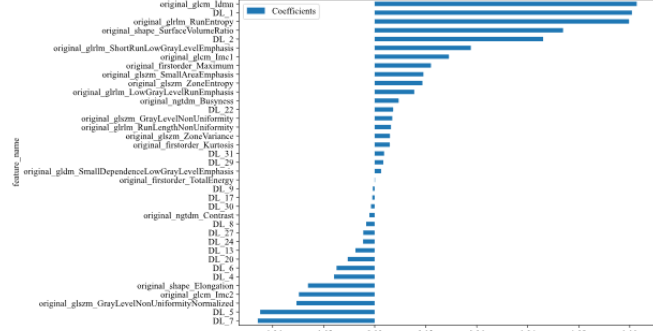
Rad



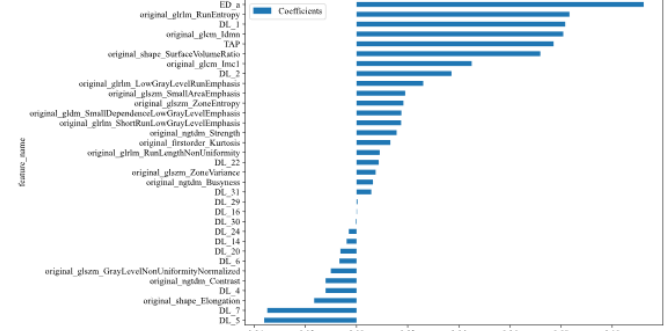
DL



DLR



Nomogram



DL, deep learning; DLR, deep learning radiomics.

## Appendix 4

Performance of radiomics features in the training and test cohort under different machine learning algorithms.

Model name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Threshold	Task
LR	0.82	0.857	0.8244–0.8902	0.885	0.703	0.842	0.774	0.534	Label-train
	0.833	0.9	0.8576–0.9424	0.839	0.826	0.874	0.78	0.598	Label-test
NaiveBayes	0.701	0.748	0.7069–0.7881	0.727	0.656	0.790	0.573	0.009	Label-train
	0.767	0.773	0.7069–0.8393	0.774	0.756	0.821	0.699	0.275	Label-test
SVM	0.816	0.893	0.8651–0.9214	0.794	0.856	0.908	0.699	0.755	Label-train
	0.724	0.807	0.7469–0.8673	0.661	0.814	0.837	0.625	0.752	Label-test
KNN	0.629	0.889	0.8638–0.9135	0.421	1.000	1.000	0.492	0.8	Label-train
	0.586	0.657	0.5824–0.7312	0.556	0.628	0.683	0.495	0.6	Label-test
LightGBM	0.833	0.897	0.8699–0.9231	0.826	0.847	0.906	0.731	0.657	Label-train
	0.729	0.765	0.6999–0.8298	0.79	0.640	0.760	0.679	0.631	Label-test
GradientBoosting	0.859	0.908	0.8841–0.9324	0.938	0.718	0.856	0.867	0.588	Label-train
	0.714	0.757	0.6905–0.8237	0.742	0.674	0.767	0.644	0.614	Label-test
AdaBoost	0.778	0.854	0.8224–0.8849	0.796	0.746	0.849	0.672	0.52	Label-train
	0.586	0.711	0.6378–0.7837	0.516	0.686	0.703	0.496	0.504	Label-test
MLP	0.789	0.857	0.8252–0.8895	0.794	0.780	0.865	0.679	0.651	Label-train
	0.805	0.854	0.8015–0.9063	0.798	0.814	0.861	0.737	0.671	Label-test

AUC, area under the curve; CI, confidence interval; PPV, **positive predictive value**; NPV, **negative predictive value**; LR, logistic regression; SVM, support vector machine; KNN, k-nearest neighbors; LightGBM, light gradient boosting machine; GradientBoosting, gradient boosting machine; AdaBoost, adaptive boosting; MLP, multi-layer perceptron.

## Appendix 5

Performance of DL features in the training and test cohort under different machine learning algorithms.

Model name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Threshold	Task
LR	0.758	0.788	0.7484–0.8269	0.807	0.670	0.814	0.66	0.582	Label-train
	0.79	0.833	0.7780–0.8883	0.782	0.802	0.851	0.719	0.619	Label-test
NaiveBayes	0.72	0.726	0.6823–0.7701	0.812	0.555	0.765	0.624	0.231	Label-train
	0.705	0.755	0.6905–0.8195	0.79	0.581	0.731	0.658	0.276	Label-test
SVM	0.799	0.858	0.8252–0.8917	0.775	0.842	0.898	0.677	0.74	Label-train
	0.767	0.761	0.6898–0.8314	0.79	0.733	0.810	0.708	0.661	Label-test
KNN	0.562	0.858	0.8286–0.8864	0.316	1.000	1.000	0.45	0.8	Label-train
	0.648	0.759	0.6945–0.8231	0.524	0.826	0.812	0.546	0.6	Label-test
LightGBM	0.84	0.929	0.9100–0.9488	0.788	0.933	0.955	0.712	0.684	Label-train
	0.748	0.758	0.6911–0.8252	0.79	0.686	0.784	0.694	0.667	Label-test
GradientBoosting	0.816	0.858	0.8262–0.8906	0.895	0.675	0.831	0.783	0.594	Label-train
	0.786	0.834	0.7786–0.8890	0.758	0.826	0.862	0.703	0.689	Label-test
AdaBoost	0.675	0.824	0.7904–0.8582	0.563	0.876	0.890	0.529	0.523	Label-train
	0.762	0.841	0.7879–0.8940	0.718	0.826	0.856	0.67	0.519	Label-test
MLP	0.78	0.805	0.7671–0.8423	0.85	0.656	0.815	0.71	0.572	Label-train
	0.786	0.819	0.7611–0.8770	0.774	0.802	0.850	0.711	0.631	Label-test

DL, deep learning; AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; KNN, k-nearest neighbors; LightGBM, light gradient boosting machine; GradientBoosting, gradient boosting machine; AdaBoost, adaptive boosting; MLP, multi-layer perceptron.

## Appendix 6

Performance of DLR features in the training and test cohort under different machine learning algorithms.

Model name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Threshold	Task
LR	0.82	0.86	0.8275–0.8926	0.882	0.708	0.844	0.771	0.534	Label-train
	0.833	0.903	0.8613–0.9442	0.823	0.849	0.887	0.768	0.6119	Label-test
NaiveBayes	0.71	0.752	0.7112–0.7929	0.769	0.603	0.776	0.594	0.0006207	Label-train
	0.748	0.807	0.7472–0.8672	0.677	0.849	0.866	0.646	0.2356	Label-test
SVM	0.833	0.917	0.8897–0.9436	0.769	0.947	0.963	0.697	0.8058	Label-train
	0.833	0.888	0.8408–0.9344	0.895	0.744	0.835	0.831	0.6334	Label-test
KNN	0.598	0.901	0.8771–0.9244	0.373	1.000	1.000	0.472	0.8	Label-train
	0.567	0.819	0.7618–0.8772	0.306	0.942	0.884	0.485	0.8	Label-test
LightGBM	0.909	0.974	0.9635–0.9836	0.877	0.967	0.979	0.815	0.6593	Label-train
	0.762	0.827	0.7702–0.8845	0.758	0.767	0.825	0.687	0.6998	Label-test
GradientBoosting	0.854	0.892	0.8626–0.9207	0.893	0.785	0.881	0.804	0.6346	Label-train
	0.781	0.828	0.7687–0.8881	0.75	0.826	0.861	0.696	0.6877	Label-test
AdaBoost	0.753	0.856	0.8257–0.8865	0.716	0.818	0.875	0.617	0.526	Label-train
	0.781	0.82	0.7613–0.8785	0.718	0.872	0.890	0.682	0.5269	Label-test
MLP	0.857	0.903	0.8779–0.9288	0.909	0.766	0.874	0.825	0.577	Label-train
	0.829	0.878	0.8305–0.9257	0.895	0.733	0.828	0.829	0.5479	Label-test

DLR, deep learning radiomics; AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; KNN, k-nearest neighbors; LightGBM, light gradient boosting machine; GradientBoosting, gradient boosting machine; AdaBoost, adaptive boosting; MLP, multi-layer perceptron.

## Appendix 7

Performance of clinical features in the training and test cohort under different machine learning algorithms.

Model name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Threshold	Task
LR	0.744	0.792	0.7545–0.8302	0.751	0.732	0.833	0.622	0.613	Label-train
	0.733	0.803	0.7447–0.8615	0.637	0.872	0.878	0.625	0.736	Label-test
NaiveBayes	0.739	0.788	0.7498–0.8259	0.761	0.699	0.818	0.621	0.579	Label-train
	0.729	0.801	0.7411–0.8600	0.629	0.872	0.876	0.62	0.751	Label-test
SVM	0.746	0.788	0.7494–0.8264	0.756	0.727	0.832	0.626	0.68	Label-train
	0.748	0.797	0.7368–0.8572	0.677	0.849	0.866	0.646	0.751	Label-test
KNN	0.722	0.848	0.8183–0.8784	0.651	0.847	0.884	0.577	0.6	Label-train
	0.619	0.778	0.7159–0.8402	0.403	0.930	0.893	0.519	0.8	Label-test
LightGBM	0.823	0.904	0.8797–0.9273	0.858	0.761	0.865	0.75	0.581	Label-train
	0.805	0.822	0.7618–0.8824	0.855	0.733	0.822	0.778	0.613	Label-test
GradientBoosting	0.773	0.832	0.7978–0.8666	0.804	0.718	0.836	0.673	0.643	Label-train
	0.729	0.813	0.7551–0.8706	0.621	0.884	0.885	0.618	0.754	Label-test
AdaBoost	0.679	0.817	0.7828–0.8510	0.563	0.885	0.897	0.532	0.518	Label-train
	0.69	0.797	0.7372–0.8563	0.573	0.860	0.855	0.583	0.527	Label-test
MLP	0.716	0.781	0.7423–0.8192	0.678	0.785	0.849	0.577	0.613	Label-train
	0.71	0.811	0.7535–0.8694	0.581	0.895	0.889	0.597	0.723	Label-test

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; KNN, k-nearest neighbors; LightGBM, light gradient boosting machine; GradientBoosting, gradient boosting machine; AdaBoost, adaptive boosting; MLP, multi-layer perceptron.