Supplementary Material

Kim Melanie Kraus1,2,3, Maksym Oreshko1,4, Julia Anne Schnabel5,6,7, Denise Bernhardt1,3, Stephanie Elisabeth Combs 1,2,3, Jan Caspar Peeken1,2,3

1Department of Radiation Oncology, School of Medicine and Klinikum rechts der Isar, Technical University of Munich (TUM), 81675 Munich, Germany

2Institute of Radiation Medicine (IRM), Helmholtz Zentrum München (HMGU) GmbH, German Research Center for Environmental Health, 85764 Neuherberg, Germany

3Partner Site Munich, German Consortium for Translational Cancer Research (DKTK), 80336 Munich, Germany

4 Medical Faculty, University hospital, LMU Munich, 80539 Munich, Germany

5School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom

6School of Computation, Information and Technology, Technical University of Munich, Germany

7Institute of Machine Learning in Biomedical Imaging, Helmholtz Zentrum München (HMGU) GmbH, German Research Center for Environmental Health, 85764 Neuherberg, Germany.

**\* Correspondence:**Kim Melanie Kraus
kimmelanie.kraus@mri.tum.de

# Supplementary Table 1: All features extracted from dose distributions or CT image data for each segment (PTV+2cm, total lung-GTV, ipsilateral lung - GTV)

|  |  |
| --- | --- |
| 1 | shape\_Elongation |
| 2 | shape\_Flatness |
| 3 | shape\_LeastAxisLength |
| 4 | shape\_MajorAxisLength |
| 5 | shape\_Maximum2DDiameterColumn |
| 6 | shape\_Maximum2DDiameterRow |
| 7 | shape\_Maximum2DDiameterSlice |
| 8 | shape\_Maximum3DDiameter |
| 9 | shape\_MeshVolume |
| 10 | shape\_MinorAxisLength |
| 11 | shape\_Sphericity |
| 12 | shape\_SurfaceArea |
| 13 | shape\_SurfaceVolumeRatio |
| 14 | shape\_VoxelVolume |
| 15 | firstorder\_Energy |
| 16 | firstorder\_Entropy |
| 17 | firstorder\_Minimum |
| 18 | firstorder\_10Percentile |
| 19 | firstorder\_90Percentile |
| 20 | firstorder\_Maximum |
| 21 | firstorder\_Mean |
| 22 | firstorder\_Median |
| 23 | firstorder\_InterquartileRange |
| 24 | firstorder\_Range |
| 25 | firstorder\_MeanAbsoluteDeviation |
| 26 | firstorder\_RootMeanSquared |
| 27 | firstorder\_Skewness |
| 28 | firstorder\_Kurtosis |
| 29 | firstorder\_Variance |
| 30 | firstorder\_Uniformity |
| 31 | glcm\_Autocorrelation |
| 32 | glcm\_JointAverage |
| 33 | glcm\_ClusterProminence |
| 34 | glcm\_ClusterShade |
| 35 | glcm\_ClusterTendency |
| 36 | glcm\_Contrast |
| 37 | glcm\_Correlation |
| 38 | glcm\_DifferenceAverage |
| 39 | glcm\_DifferenceEntropy |
| 40 | glcm\_DifferenceVariance |
| 41 | glcm\_JointEnergy |
| 42 | glcm\_JointEntropy |
| 43 | glcm\_Imc1 |
| 44 | glcm\_Imc2 |
| 45 | glcm\_Idm |
| 46 | glcm\_Idmn |
| 47 | glcm\_MCC |
| 48 | glcm\_Id |
| 49 | glcm\_Idn |
| 50 | glcm\_InverseVariance |
| 51 | glcm\_MaximumProbability |
| 52 | glcm\_SumEntropy |
| 53 | glcm\_SumSquares |
| 54 | glrlm\_GrayLevelNonUniformity |
| 55 | glrlm\_GrayLevelNonUniformityNormalized |
| 56 | glrlm\_GrayLevelVariance |
| 57 | glrlm\_HighGrayLevelRunEmphasis |
| 58 | glrlm\_LongRunEmphasis |
| 59 | glrlm\_LongRunHighGrayLevelEmphasis |
| 60 | glrlm\_LongRunLowGrayLevelEmphasis |
| 61 | glrlm\_LowGrayLevelRunEmphasis |
| 62 | glrlm\_RunEntropy |
| 63 | glrlm\_RunLengthNonUniformity |
| 64 | glrlm\_RunLengthNonUniformityNormalized |
| 65 | glrlm\_RunPercentage |
| 66 | glrlm\_RunVariance |
| 67 | glrlm\_ShortRunEmphasis |
| 68 | glrlm\_ShortRunHighGrayLevelEmphasis |
| 69 | glrlm\_ShortRunLowGrayLevelEmphasis |
| 70 | glszm\_GrayLevelNonUniformity |
| 71 | glszm\_GrayLevelNonUniformityNormalized |
| 72 | glszm\_GrayLevelVariance |
| 73 | glszm\_HighGrayLevelZoneEmphasis |
| 74 | glszm\_LargeAreaEmphasis |
| 75 | glszm\_LargeAreaHighGrayLevelEmphasis |
| 76 | glszm\_LargeAreaLowGrayLevelEmphasis |
| 77 | glszm\_LowGrayLevelZoneEmphasis |
| 78 | glszm\_SizeZoneNonUniformity |
| 79 | glszm\_SizeZoneNonUniformityNormalized |
| 80 | glszm\_SmallAreaEmphasis |
| 81 | glszm\_SmallAreaHighGrayLevelEmphasis |
| 82 | glszm\_SmallAreaLowGrayLevelEmphasis |
| 83 | glszm\_ZoneEntropy |
| 84 | glszm\_ZonePercentage |
| 85 | glszm\_ZoneVariance |
| 86 | ngtdm\_Busyness |
| 87 | ngtdm\_Coarseness |
| 88 | ngtdm\_Complexity |
| 89 | ngtdm\_Contrast |
| 90 | ngtdm\_Strength |
| 91 | gldm\_DependenceEntropy |
| 92 | gldm\_DependenceNonUniformity |
| 93 | gldm\_DependenceNonUniformityNormalized |
| 94 | gldm\_DependenceVariance |
| 95 | gldm\_GrayLevelNonUniformity |
| 96 | gldm\_GrayLevelVariance |
| 97 | gldm\_HighGrayLevelEmphasis |
| 98 | gldm\_LargeDependenceEmphasis |
| 99 | gldm\_LargeDependenceHighGrayLevelEmphasis |
| 100 | gldm\_LargeDependenceLowGrayLevelEmphasis |
| 101 | gldm\_LowGrayLevelEmphasis |
| 102 | gldm\_SmallDependenceEmphasis |
| 103 | gldm\_SmallDependenceHighGrayLevelEmphasis |
| 104 | gldm\_SmallDependenceLowGrayLevelEmphasis |

# Supplementary Table 2: all extracted features after feature reduction used in this study. Abbreviations used: planning target volume (PTV), total lung-GTV (TL), ipsilateral lung-GTV (IL)

|  |  |  |
| --- | --- | --- |
| Features | EQD2 | Physical Dose |
| Clinics | Age | Age |
|  | Size | Size |
|  | Location | Location |
|  | Sex | Sex |
|  | KPI | KPI |
| Dosimetric Features | DVH\_V50 | DVH\_V50 |
| Dosimetric Features & Clinics | DVH\_V50 | DVH\_V50 |
|  | Age | Age |
|  | Size | Size |
|  | Location | Location |
|  | Sex | Sex |
|  | KPI | KPI |
| Dosiomics | PTV\_original\_glszm\_GrayLevelNonUniformity | TL\_original\_firstorder\_InterquartileRange |
|  | PTV\_original\_glrlm\_RunEntropy | PTV\_original\_gldm\_DependenceNonUniformityNormalized |
|  | IL\_original\_shape\_Sphericity | TL\_original\_shape\_Sphericity |
|  | TL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | IL\_original\_shape\_Sphericity |
|  | TL\_original\_shape\_Sphericity | IL\_original\_shape\_LeastAxisLength |
|  | IL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | PTV\_original\_firstorder\_Minimum |
|  | TL\_original\_shape\_SurfaceVolumeRatio | TL\_original\_shape\_Flatness |
|  | PTV\_original\_glszm\_GrayLevelNonUniformityNormalized | PTV\_original\_shape\_Elongation |
|  | IL\_original\_shape\_LeastAxisLength | TL\_original\_glrlm\_RunEntropy |
|  | IL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | IL\_original\_shape\_Elongation |
|  | TL\_original\_shape\_Flatness | TL\_original\_shape\_Elongation |
|  | TL\_original\_firstorder\_InterquartileRange | PTV\_original\_gldm\_DependenceVariance |
|  | TL\_original\_firstorder\_Minimum | TL\_original\_firstorder\_Minimum |
|  | TL\_original\_shape\_Elongation | PTV\_original\_glszm\_ZoneEntropy |
|  | IL\_original\_glcm\_Imc1 | IL\_original\_glrlm\_RunEntropy |
|  | TL\_original\_glszm\_LowGrayLevelZoneEmphasis | TL\_original\_shape\_SurfaceVolumeRatio |
|  | IL\_original\_shape\_Elongation | IL\_original\_shape\_SurfaceVolumeRatio |
|  | PTV\_original\_glcm\_MCC | PTV\_original\_glszm\_SmallAreaLowGrayLevelEmphasis |
|  | IL\_original\_glszm\_GrayLevelNonUniformityNormalized | IL\_original\_glcm\_Idmn |
|  | TL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis |  |
|  | IL\_original\_glrlm\_RunEntropy |  |
| Radiomics | CT\_PTV\_original\_ngtdm\_Strength | CT\_PTV\_original\_ngtdm\_Strength |
|  | CT\_IL\_original\_shape\_Sphericity | CT\_IL\_original\_shape\_Sphericity |
|  | CT\_IL\_original\_glcm\_Idn | CT\_IL\_original\_glcm\_Idn |
|  | CT\_GL\_original\_shape\_Sphericity | CT\_GL\_original\_shape\_Sphericity |
|  | CT\_PTV\_original\_glcm\_ClusterProminence | CT\_PTV\_original\_glcm\_ClusterProminence |
|  | CT\_IL\_original\_ngtdm\_Contrast | CT\_IL\_original\_ngtdm\_Contrast |
|  | CT\_PTV\_original\_firstorder\_Energy | CT\_PTV\_original\_firstorder\_Energy |
|  | CT\_PTV\_original\_glrlm\_RunEntropy | CT\_IL\_original\_shape\_Elongation |
|  | CT\_PTV\_original\_ngtdm\_Complexity | CT\_PTV\_original\_ngtdm\_Complexity |
|  | CT\_IL\_original\_shape\_Elongation | CT\_PTV\_original\_glrlm\_RunEntropy |
|  | CT\_PTV\_original\_glcm\_ClusterShade | CT\_PTV\_original\_glcm\_ClusterShade |
|  | CT\_GL\_original\_glcm\_MCC | CT\_GL\_original\_glcm\_MCC |
|  | CT\_GL\_original\_ngtdm\_Contrast | CT\_GL\_original\_ngtdm\_Contrast |
|  | CT\_IL\_original\_glcm\_ClusterShade | CT\_IL\_original\_glcm\_ClusterShade |
|  | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis |
|  | CT\_IL\_original\_shape\_Flatness | CT\_IL\_original\_shape\_Flatness |
|  | CT\_IL\_original\_shape\_LeastAxisLength | CT\_IL\_original\_shape\_LeastAxisLength |
|  | CT\_GL\_original\_shape\_SurfaceVolumeRatio | CT\_GL\_original\_shape\_SurfaceVolumeRatio |
|  | CT\_PTV\_original\_glrlm\_GrayLevelVariance | CT\_PTV\_original\_glrlm\_GrayLevelVariance |
|  | CT\_PTV\_original\_glcm\_InverseVariance | CT\_PTV\_original\_glcm\_InverseVariance |
|  | CT\_GL\_original\_shape\_Elongation | CT\_IL\_original\_ngtdm\_Strength |
|  | CT\_IL\_original\_ngtdm\_Strength | CT\_GL\_original\_shape\_Elongation |
|  | CT\_GL\_original\_shape\_Flatness | CT\_GL\_original\_shape\_Flatness |
| Dosiomics+Radiomics | CT\_PTV\_original\_ngtdm\_Strength | CT\_PTV\_original\_ngtdm\_Strength |
|  | PTV\_original\_glszm\_GrayLevelNonUniformity | CT\_IL\_original\_glcm\_Idn |
|  | CT\_IL\_original\_glcm\_Idn | CT\_PTV\_original\_glcm\_ClusterProminence |
|  | CT\_PTV\_original\_glcm\_ClusterProminence | TL\_original\_firstorder\_InterquartileRange |
|  | TL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | CT\_IL\_original\_ngtdm\_Contrast |
|  | CT\_IL\_original\_ngtdm\_Contrast | PTV\_original\_gldm\_DependenceNonUniformityNormalized |
|  | PTV\_original\_glrlm\_RunEntropy | CT\_PTV\_original\_firstorder\_Energy |
|  | CT\_PTV\_original\_firstorder\_Energy | CT\_PTV\_original\_ngtdm\_Complexity |
|  | CT\_PTV\_original\_ngtdm\_Complexity | CT\_PTV\_original\_glrlm\_RunEntropy |
|  | IL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | CT\_GL\_original\_shape\_Sphericity |
|  | CT\_PTV\_original\_glrlm\_RunEntropy | PTV\_original\_firstorder\_Minimum |
|  | PTV\_original\_glszm\_GrayLevelNonUniformityNormalized | CT\_PTV\_original\_glcm\_ClusterShade |
|  | CT\_PTV\_original\_glcm\_ClusterShade | CT\_GL\_original\_glcm\_MCC |
|  | CT\_GL\_original\_glcm\_MCC | TL\_original\_glrlm\_RunEntropy |
|  | IL\_original\_shape\_Sphericity | IL\_original\_shape\_Sphericity |
|  | TL\_original\_glszm\_LowGrayLevelZoneEmphasis | CT\_GL\_original\_ngtdm\_Contrast |
|  | IL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | PTV\_original\_shape\_Elongation |
|  | CT\_GL\_original\_ngtdm\_Contrast | CT\_IL\_original\_glcm\_ClusterShade |
|  | TL\_original\_firstorder\_InterquartileRange | CT\_IL\_original\_shape\_Sphericity |
|  | CT\_GL\_original\_shape\_Sphericity | CT\_PTV\_original\_glrlm\_GrayLevelVariance |
|  | CT\_PTV\_original\_glrlm\_GrayLevelVariance | CT\_PTV\_original\_glcm\_InverseVariance |
|  | CT\_IL\_original\_glcm\_ClusterShade | IL\_original\_shape\_LeastAxisLength |
|  | IL\_original\_glcm\_Imc1 | TL\_original\_firstorder\_Minimum |
|  | TL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | IL\_original\_glrlm\_RunEntropy |
|  | CT\_PTV\_original\_glcm\_InverseVariance | PTV\_original\_glszm\_ZoneEntropy |
|  | TL\_original\_firstorder\_Minimum | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis |
|  | CT\_IL\_original\_shape\_Sphericity | CT\_IL\_original\_ngtdm\_Strength |
|  | IL\_original\_glszm\_GrayLevelNonUniformityNormalized | PTV\_original\_gldm\_DependenceVariance |
|  | IL\_original\_glrlm\_RunEntropy | CT\_IL\_original\_glszm\_GrayLevelNonUniformity |
|  | TL\_original\_firstorder\_Kurtosis | IL\_original\_glcm\_Idmn |
|  | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis | CT\_IL\_original\_firstorder\_Minimum |
|  | CT\_GL\_original\_shape\_SurfaceVolumeRatio | CT\_GL\_original\_shape\_SurfaceVolumeRatio |
|  | PTV\_original\_glszm\_ZoneEntropy | CT\_GL\_original\_shape\_Elongation |
|  | CT\_PTV\_original\_glszm\_LargeAreaLowGrayLevelEmphasis |  |
| RadiomicsDVHClinics | CT\_PTV\_original\_ngtdm\_Strength | CT\_PTV\_original\_ngtdm\_Strength |
|  | CT\_IL\_original\_shape\_Sphericity | CT\_IL\_original\_shape\_Sphericity |
|  | CT\_IL\_original\_glcm\_Idn | CT\_IL\_original\_glcm\_Idn |
|  | CT\_GL\_original\_shape\_Sphericity | CT\_GL\_original\_shape\_Sphericity |
|  | CT\_PTV\_original\_glcm\_ClusterProminence | CT\_PTV\_original\_glcm\_ClusterProminence |
|  | CT\_IL\_original\_ngtdm\_Contrast | CT\_IL\_original\_ngtdm\_Contrast |
|  | CT\_PTV\_original\_firstorder\_Energy | CT\_PTV\_original\_firstorder\_Energy |
|  | CT\_PTV\_original\_glrlm\_RunEntropy | CT\_IL\_original\_shape\_Elongation |
|  | CT\_PTV\_original\_ngtdm\_Complexity | CT\_PTV\_original\_ngtdm\_Complexity |
|  | CT\_IL\_original\_shape\_Elongation | CT\_PTV\_original\_glrlm\_RunEntropy |
|  | CT\_PTV\_original\_glcm\_ClusterShade | CT\_PTV\_original\_glcm\_ClusterShade |
|  | CT\_GL\_original\_glcm\_MCC | CT\_GL\_original\_glcm\_MCC |
|  | CT\_GL\_original\_ngtdm\_Contrast | CT\_GL\_original\_ngtdm\_Contrast |
|  | CT\_IL\_original\_glcm\_ClusterShade | CT\_IL\_original\_glcm\_ClusterShade |
|  | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis |
|  | CT\_IL\_original\_shape\_Flatness | CT\_IL\_original\_shape\_Flatness |
|  | CT\_IL\_original\_shape\_LeastAxisLength | CT\_IL\_original\_shape\_LeastAxisLength |
|  | CT\_GL\_original\_shape\_SurfaceVolumeRatio | CT\_GL\_original\_shape\_SurfaceVolumeRatio |
|  | CT\_PTV\_original\_glrlm\_GrayLevelVariance | CT\_PTV\_original\_glrlm\_GrayLevelVariance |
|  | CT\_PTV\_original\_glcm\_InverseVariance | CT\_PTV\_original\_glcm\_InverseVariance |
|  | CT\_GL\_original\_shape\_Elongation | CT\_IL\_original\_ngtdm\_Strength |
|  | CT\_IL\_original\_ngtdm\_Strength | CT\_GL\_original\_shape\_Elongation |
|  | CT\_GL\_original\_shape\_Flatness | CT\_GL\_original\_shape\_Flatness |
|  | DVH\_V50 | DVH\_V50 |
|  | Age | Age |
|  | Size | Size |
|  | Location | Location |
|  | Sex | Sex |
|  | KPI | KPI |
| DosiomicsDVHClinics | PTV\_original\_glszm\_GrayLevelNonUniformity | TL\_original\_firstorder\_InterquartileRange |
|  | PTV\_original\_glrlm\_RunEntropy | PTV\_original\_gldm\_DependenceNonUniformityNormalized |
|  | IL\_original\_shape\_Sphericity | TL\_original\_shape\_Sphericity |
|  | TL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | IL\_original\_shape\_Sphericity |
|  | TL\_original\_shape\_Sphericity | IL\_original\_shape\_LeastAxisLength |
|  | IL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | PTV\_original\_firstorder\_Minimum |
|  | TL\_original\_shape\_SurfaceVolumeRatio | TL\_original\_shape\_Flatness |
|  | PTV\_original\_glszm\_GrayLevelNonUniformityNormalized | PTV\_original\_shape\_Elongation |
|  | IL\_original\_shape\_LeastAxisLength | TL\_original\_glrlm\_RunEntropy |
|  | IL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | IL\_original\_shape\_Elongation |
|  | TL\_original\_shape\_Flatness | TL\_original\_shape\_Elongation |
|  | TL\_original\_firstorder\_InterquartileRange | PTV\_original\_gldm\_DependenceVariance |
|  | TL\_original\_firstorder\_Minimum | TL\_original\_firstorder\_Minimum |
|  | TL\_original\_shape\_Elongation | PTV\_original\_glszm\_ZoneEntropy |
|  | IL\_original\_glcm\_Imc1 | IL\_original\_glrlm\_RunEntropy |
|  | TL\_original\_glszm\_LowGrayLevelZoneEmphasis | TL\_original\_shape\_SurfaceVolumeRatio |
|  | IL\_original\_shape\_Elongation | IL\_original\_shape\_SurfaceVolumeRatio |
|  | PTV\_original\_glcm\_MCC | PTV\_original\_glszm\_SmallAreaLowGrayLevelEmphasis |
|  | IL\_original\_glszm\_GrayLevelNonUniformityNormalized | IL\_original\_glcm\_Idmn |
|  | TL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | DVH\_V50 |
|  | IL\_original\_glrlm\_RunEntropy | Age |
|  | DVH\_V50 | Size |
|  | Age | Location |
|  | Size | Sex |
|  | Location | KPI |
|  | Sex |  |
|  | KPI |  |
| DosiomicsRadiomicsDVHClinics | CT\_PTV\_original\_ngtdm\_Strength | CT\_PTV\_original\_ngtdm\_Strength |
|  | PTV\_original\_glszm\_GrayLevelNonUniformity | CT\_IL\_original\_glcm\_Idn |
|  | CT\_IL\_original\_glcm\_Idn | CT\_PTV\_original\_glcm\_ClusterProminence |
|  | CT\_PTV\_original\_glcm\_ClusterProminence | TL\_original\_firstorder\_InterquartileRange |
|  | TL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | CT\_IL\_original\_ngtdm\_Contrast |
|  | CT\_IL\_original\_ngtdm\_Contrast | PTV\_original\_gldm\_DependenceNonUniformityNormalized |
|  | PTV\_original\_glrlm\_RunEntropy | CT\_PTV\_original\_firstorder\_Energy |
|  | CT\_PTV\_original\_firstorder\_Energy | CT\_PTV\_original\_ngtdm\_Complexity |
|  | CT\_PTV\_original\_ngtdm\_Complexity | CT\_PTV\_original\_glrlm\_RunEntropy |
|  | IL\_original\_glszm\_SmallAreaLowGrayLevelEmphasis | CT\_GL\_original\_shape\_Sphericity |
|  | CT\_PTV\_original\_glrlm\_RunEntropy | PTV\_original\_firstorder\_Minimum |
|  | PTV\_original\_glszm\_GrayLevelNonUniformityNormalized | CT\_PTV\_original\_glcm\_ClusterShade |
|  | CT\_PTV\_original\_glcm\_ClusterShade | CT\_GL\_original\_glcm\_MCC |
|  | CT\_GL\_original\_glcm\_MCC | TL\_original\_glrlm\_RunEntropy |
|  | IL\_original\_shape\_Sphericity | IL\_original\_shape\_Sphericity |
|  | TL\_original\_glszm\_LowGrayLevelZoneEmphasis | CT\_GL\_original\_ngtdm\_Contrast |
|  | IL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | PTV\_original\_shape\_Elongation |
|  | CT\_GL\_original\_ngtdm\_Contrast | CT\_IL\_original\_glcm\_ClusterShade |
|  | TL\_original\_firstorder\_InterquartileRange | CT\_IL\_original\_shape\_Sphericity |
|  | CT\_GL\_original\_shape\_Sphericity | CT\_PTV\_original\_glrlm\_GrayLevelVariance |
|  | CT\_PTV\_original\_glrlm\_GrayLevelVariance | CT\_PTV\_original\_glcm\_InverseVariance |
|  | CT\_IL\_original\_glcm\_ClusterShade | IL\_original\_shape\_LeastAxisLength |
|  | IL\_original\_glcm\_Imc1 | TL\_original\_firstorder\_Minimum |
|  | TL\_original\_glrlm\_ShortRunLowGrayLevelEmphasis | IL\_original\_glrlm\_RunEntropy |
|  | CT\_PTV\_original\_glcm\_InverseVariance | PTV\_original\_glszm\_ZoneEntropy |
|  | TL\_original\_firstorder\_Minimum | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis |
|  | CT\_IL\_original\_shape\_Sphericity | CT\_IL\_original\_ngtdm\_Strength |
|  | IL\_original\_glszm\_GrayLevelNonUniformityNormalized | PTV\_original\_gldm\_DependenceVariance |
|  | IL\_original\_glrlm\_RunEntropy | CT\_IL\_original\_glszm\_GrayLevelNonUniformity |
|  | TL\_original\_firstorder\_Kurtosis | IL\_original\_glcm\_Idmn |
|  | CT\_GL\_original\_glszm\_LargeAreaHighGrayLevelEmphasis | CT\_IL\_original\_firstorder\_Minimum |
|  | CT\_GL\_original\_shape\_SurfaceVolumeRatio | CT\_GL\_original\_shape\_SurfaceVolumeRatio |
|  | PTV\_original\_glszm\_ZoneEntropy | CT\_GL\_original\_shape\_Elongation |
|  | CT\_PTV\_original\_glszm\_LargeAreaLowGrayLevelEmphasis | DVH\_V50 |
|  | DVH\_V50 | Age |
|  | Age | Size |
|  | Size | Location |
|  | Location | Sex |
|  | Sex | KPI |
|  | KPI |  |

# Supplementary Table 3: Hyperparameter tuning grid for machine learning models

|  |  |  |
| --- | --- | --- |
|  | **Hyperparameter** | **Search space (by increment)** |
| **Elastic Net Regression** | Lambda  | 0 – 1, by 0.02 |
| Alpha | 0 – 1, by 0.05 |
| **Random Forest** | Number of trees (ntree) | 501 – 2001, by 500 |
| Features per node (Mtry) | 1 – 6, by 1 |
| **LogitBoost** | Iteration Number | 1 – 100, by 2 |
| **Support Vector Machine** | C | 0.00001 – 0.01. by 0.005, 0.01 – 0.1, by 0.05,0.1 – 1, by 0.2,1 – 10, by 2 |
| Sigma | 0.1 – 1, by 0.1,1 –50, by 5, 60 – 450, by 70 |

# 4 Calibration Curves of the nine best performing models



Figure 1: Calibration Curve for Radiomics model.



Figure 2: Calibration Curve for Dosiomics model for physical dose.



Figure 3: Calibration model for combined Dosiomics and Radiomics model for physical dose.



Figure 4: Calibration model for combined Radiomics+DVH+Clinics model for physical dose.



Figure 5: Calibration Curve for all models combined (ALL) for physical dose.



Figure 6: Calibration curve for Dosiomics model for EQD2.



Figure 7: Calibration model for combined Dosiomics and Radiomics model for EQD2.



Figure 8: Calibration curve for Dosiomics+DVH+Clinics for EQD2.



Figure 9: Calibration curve for all models combined (ALL) for EQD2.

# 5 Supplementary Table 4: Tripod Checklist: Prediction Model Development and Validation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Section/Topic** | **Item** |  | **Checklist Item** | **Page** | **Reported on section** |
| **Title and abstract** |  |
| Title | 1 | D;V | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | 1 | Title |
| Abstract | 2 | D;V | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | 2 | Abstract |
| **Introduction** |  |
| Background and objectives | 3a | D;V | * Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.
 | 3 | Introduction |
| 3b | D;V | * Specify the objectives, including whether the study describes the development or validation of the model or both.
 | 3, last paragraph | Introduction |
| **Methods** |  |
| Source of data | 4a | D;V | * Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.
 | 3-5, Section 2.1Section 2.2 | Methods. 2.1 Patient clinical Factors2.2 Volumetric CT and Dose Data |
| 4b | D;V | * Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.
 | 3-4, Section 2.1 | Methods. Patient clinical Factors |
| Participants | 5a | D;V | * Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.
 | 3-4, Section 2.1 | Methods. Patient clinical Factors |
| 5b | D;V | * Describe eligibility criteria for participants.
 | 3-4, Section 2.1 | Methods. Patient clinical Factors |
| 5c | D;V | * Give details of treatments received, if relevant.
 | 3-5, Section 2.1 | Methods. Patient clinical Factors |
| Outcome | 6a | D;V | * Clearly define the outcome that is predicted by the prediction model, including how and when assessed.
 |  |  |
| 6b | D;V | * Report any actions to blind assessment of the outcome to be predicted.
 | - | - |
| Predictors | 7a | D;V | * Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.
 | 6-7Section 2.3 / 2.4 | Methods 2.3 Feature Processing2.4 Machine Learning Model |
| 7b | D;V | * Report any actions to blind assessment of predictors for the outcome and other predictors.
 | - | - |
| Sample size | 8 | D;V | * Explain how the study size was arrived at.
 | - | - |
| Missing data | 9 | D;V | * Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.
 | - | - |
| Statistical analysis methods | 10a | D | * Describe how predictors were handled in the analyses.
 | 7Section 2.4 | 2.4 Machine Learning Model |
| 10b | D | * Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.
 | 6-7* Section 2.3 / 2.4
 | Methods 2.3 Feature Processing* 2.4 Machine Learning Model
 |
| 10c | V | * For validation, describe how the predictions were calculated.
 | 7* Section 2.4
 | * 2.4 Machine Learning Model
 |
| 10d | D;V | * Specify all measures used to assess model performance and, if relevant, to compare multiple models.
 | 7Section 2.4 | 2.4 Machine Learning Model |
| 10e | V | * Describe any model updating (e.g., recalibration) arising from the validation, if done.
 | 7Section 2.4 | 2.4 Machine Learning Model |
| Risk groups | 11 | D;V | Provide details on how risk groups were created, if done.  |  | - |
| Development vs. validation | 12 | V | For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.  | - | - |
| **Results** |  |
| Participants | 13a | D;V | * Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.
 | 2.4 Machine Learning Model | Methods. Patient clinical Factors |
| 13b | D;V | * Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.
 | 2.4 Machine Learning Model | Methods. Patient clinical Factors |
| 13c | V | * For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).
 | * -
 | * -
 |
| Model development  | 14a | D | * Specify the number of participants and outcome events in each analysis.
 | * 8-9
 | * Table 2
 |
| 14b | D | * If done, report the unadjusted association between each candidate predictor and outcome.
 |  |  |
| Model specification | 15a | D | * Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).
 |  |  |
| 15b | D | * Explain how to the use the prediction model.
 |  |  |
| Model performance | 16 | D;V | * Report performance measures (with CIs) for the prediction model.
 | 8-10 | Results3.1 Classifiers3.3 Machine Learning PTP Prediction |
| Model-updating | 17 | V | If done, report the results from any model updating (i.e., model specification, model performance). |  |  |
| **Discussion** |  |
| Limitations | 18 | D;V | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).  | 11 | Discussion, paragraph 6 & 7 |
| Interpretation | 19a | V | * For validation, discuss the results with reference to performance in the development data, and any other validation data.
 | * 10-11
 | * Discussion, paragraphs 2,3,4 & 5
 |
| 19b | D;V | * Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.
 | 11 | Conclusions |
| Implications | 20 | D;V | Discuss the potential clinical use of the model and implications for future research.  |  |  |
| **Other information** |  |
| Supplementary information | 21 | D;V | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.  | 12 | 9 Data Availability Statement |
| Funding | 22 | D;V | Give the source of funding and the role of the funders for the present study.  | 12 | 8 Funding |

# 6 Supplementary Data 6: AUC values for all Classifiers and all models with and without shape features



Figure 10: AUC values for all classifiers and models for physical dose (D) and 2-Gy-equivalent dose (EQD2). For individual models, shape features were excluded from the simulation.