

Supplementary: Statistical method

Recursive feature elimination (RFE)

As a well-known and commonly used feature selection algorithm, Recursive Feature Elimination (RFE), a backward selection approach with the utilization of stepwise method [1-2], implements backwards selection of predictors according to predictor importance ranking. The survival model, as random survival forest (R package *mlr3extralearners*: `lrn("surv.ranger")`), with the right censored time was first trained on all potential predictors, the importance score was obtained via the permutation method. The predictor with lowest importance score was excluded. The model was refitted with the reduced pool of predictors as next step, following the model evaluation with 5-fold cross validation approach with C-index as performance measure. These steps were repeated iteratively until all features were traversed. This process enables the identification of optimal collection of predictors that this collection of predictors will be included into our final model [3].

Tuning process and hyperparameter of sSVM [4]

Two hyperparameters of sSVM, gamma and mu, was tuned with the following settings (R package *mlr3tuning*: `AutoTuner()`):

1. The hyperparameters of the wrapped (inner) learner are trained on the training data via resampling, which was specified as 5-folded cross-validation.
2. The measure to be optimized is set to be C-index.
3. Turner is set to apply the method of grid search.
4. Hyperparameter search space was set to be (0.001 and 1000) for both parameters gamma and mu.
5. The termination criteria was set to be after 4 evaluations

Supplementary Table S1: potential dosimetric and clinical variables to be selected from

	Dosimetric variables	Clinical variables
1	PTV_V100	Histology.in.Number..Epithelioid.0..Mixed..1.
2	PTV_Volume.cc.	gender_male
3	PTV_Max	T.Stage
4	PTV_Mean	N.Stage
5	PTV_Min	age
6	PTV.ipsi.PTV	Pneumonitis
7	Total_Lung_PTV_Volume	PTV_Side..L
8	Total_Lung_PTV_Mean	
9	Total_Lung_PTV_V20	
10	Ipsi_Lung_PTV_Volume	
11	Ipsi_Lung_PTV_Mean	

12	Ipsi_Lung_PTV_V20	
13	Contra_Lung_Volume	
14	Contra_Lung_Mean	
15	Contra_Lung_V5	
16	Contra_Lung_V20	
17	Esophagus_Volume	
18	Esophagus_Max	
19	Esophagus_Mean	
20	Heart_Volume	
21	Heart_Mean	
22	Heart_V5	
23	Heart_V30	
24	Heart_V50	
25	Spinal_Cord_Volume	
26	Spinal_Cord_Max	

Supplementary Table S2: summary of model performance for different settings

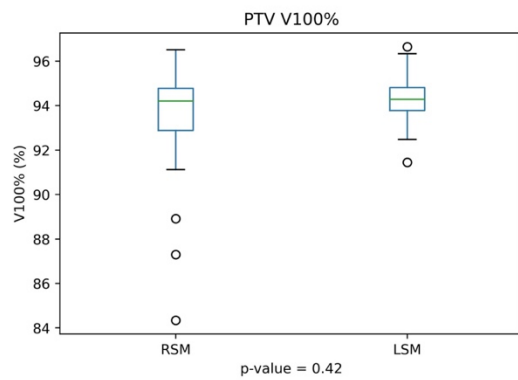
	Metrics of predictive performance**						
Features	C-index (5-fold cross validation)	van Houwelingen's Alpha Calibration	van Houwelingen's Beta Calibration	Integrated Graf Score	Integrated Log Loss	Log Loss	Hyperparameter (tuned)
Dosimetric variable – right							
Cox PH	0.54	0.69	0.02	0.27	1.10	34.54	
sSVM	0.56	NA	NA	NA	NA	NA	Gamma=1000, mu=0.001
Dosimetric variable – left							
Cox PH	0.47	Inf	-0.00	0.54	3.71	34.54	
sSVM	0.41	NA	NA	NA	NA	NA	Gamma=1000, mu=1000
Clinical variable							
Cox PH	0.70	1.12	3.29	0.21	0.62	31.39	
sSVM	0.65	NA	NA	NA	NA	NA	Gamma=0.001, mu=1000

Combined dosimetric and clinical variables							
Cox PH	0.62	0.36	0.26	0.29	1.14	31.18	
sSVM	0.64	NA	NA	NA	NA	NA	Gamma=0.001, mu=1000

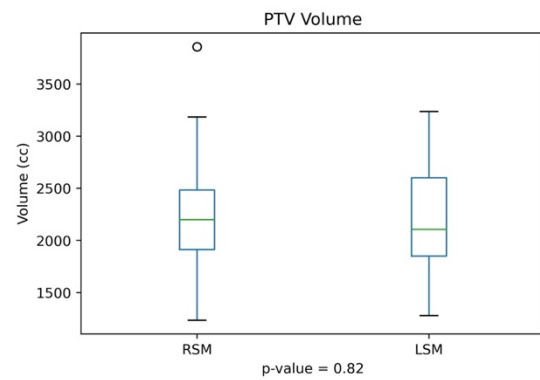
*Gray indicates that stable fit was unattainable.

**the metrics of predictive performance are obtained via R package mlr3proba [5].

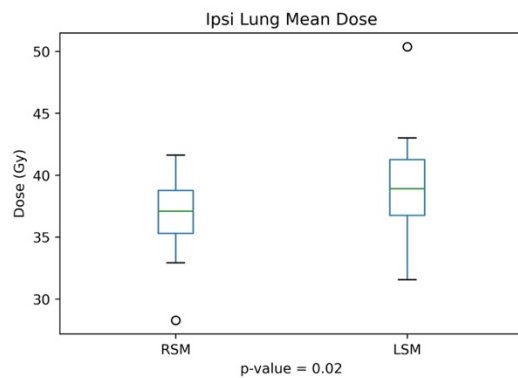
Supplemental Figure S1: Box plots of the selected dosimetric metrics between LSM and RSM groups. The dosimetric metrics include (a) PTV V100%, (b) PTV volume, (c) ipsilateral lung mean dose, (d) ipsilateral lung V20Gy, I ipsilateral lu-g - PTV mean dose, (f) contralateral kidney mean dose, (g) heart mean dose, (h) liver mean dose, (i) liver V30Gy, and (j) stomach mean dose.



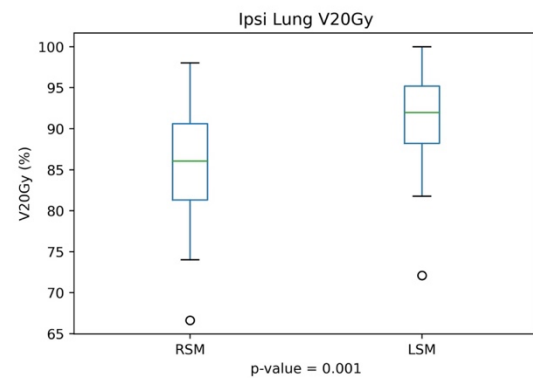
(a)



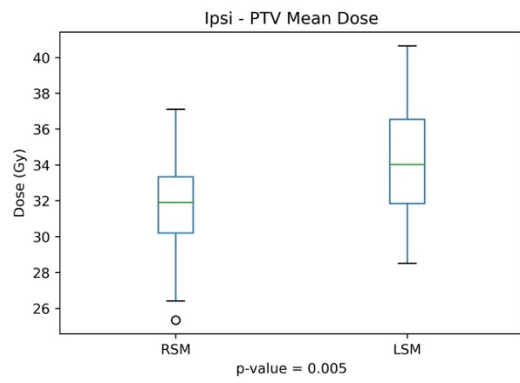
(b)



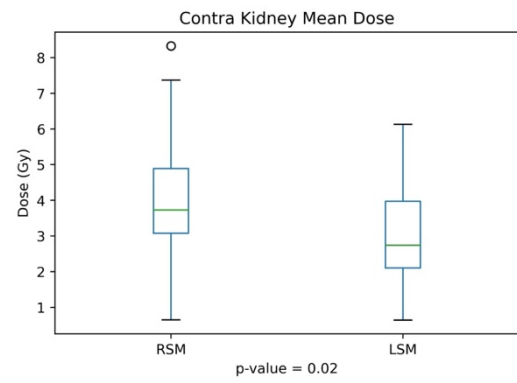
(c)



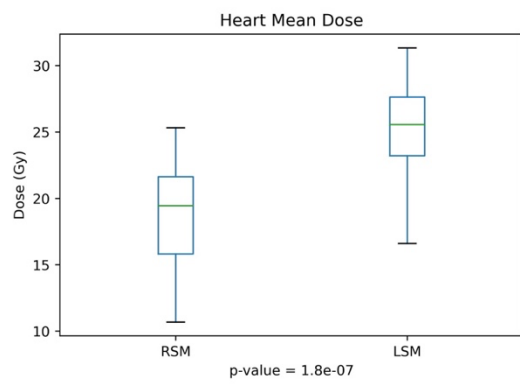
(d)



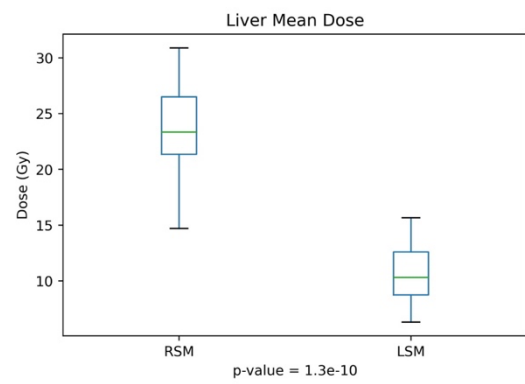
(e)



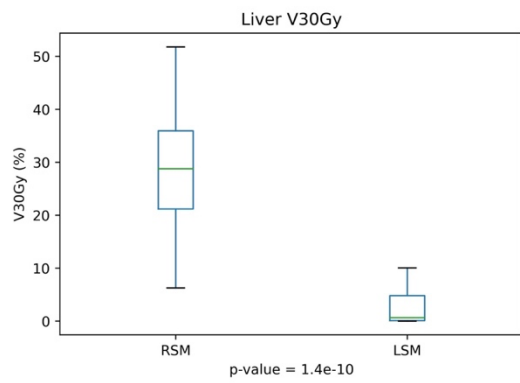
(f)



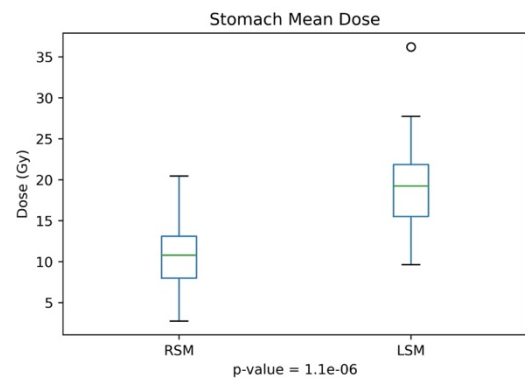
(g)



(h)



(i)



(j)

Reference:

1. Escanilla NS, Hellerstein L, Kleiman R, Kuang Z, Shull JD, Page D. Recursive Feature Elimination by Sensitivity Testing. *Proc Int Conf Mach Learn Appl*. 2018 Dec;2018:40-47. doi: 10.1109/ICMLA.2018.00014. Epub 2019 Jan 17. PMID: 31799516; PMCID: PMC6887481.
2. Giorgos Borboudakis and Ioannis Tsamardinos. 2019. Forward-backward selection with early dropping. *J. Mach. Learn. Res.* 20, 1 (January 2019), 276–314
3. Wu YL, Li WF, Yang KB, Chen L, Shi JR, Chen FP, Huang XD, Lin L, Zhang XM, Li J, Chen YP, Tang LL, Mao YP, Ma J. Long-Term Evaluation and Normal Tissue Complication Probability (NTCP) Models for Predicting Radiation-Induced Optic Neuropathy after Intensity-Modulated Radiation Therapy (IMRT) for Nasopharyngeal Carcinoma: A Large Retrospective Study in China. *J Oncol*. 2022 Feb 23;2022:3647462. doi: 10.1155/2022/3647462. PMID: 35251172; PMCID: PMC8890878.
4. Marc Becker, Michel Lang, Jakob Richter, Bernd Bischl and Daniel Schalk (2022). *mlr3tuning: Tuning for 'mlr3'*. R package version 0.15.0. <https://CRAN.R-project.org/package=mlr3tuning>
5. Sonabend R, Király FJ, Bender A, Bischl B, Lang M (2021). “*mlr3proba: An R Package for Machine Learning in Survival Analysis.*” *Bioinformatics*. ISSN 1367-4803, doi:10.1093/bioinformatics/btab039

Correlation Matrix of the Predictors

[illegible]