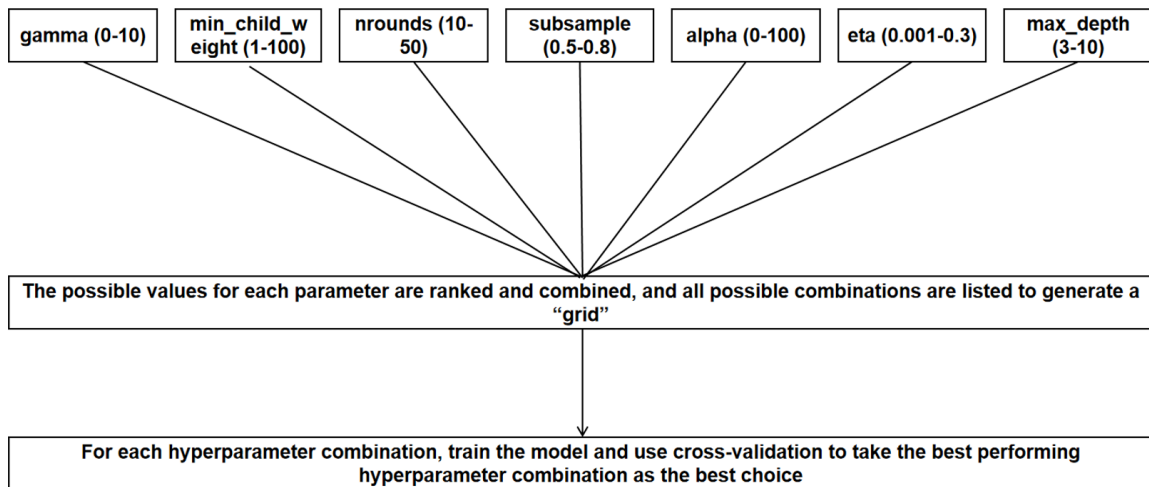


# Machine learning-based prognostic models of lung cancer brain metastases

**Supplementary Table 1.** Main parameters of the XGBoost model

Parameter	Value
Gamma	1
Scale_pos_weight	1
Min_child_weight	5
Subsample	0.7
Alpha	1
Nround	20
Max_depth	5
Eta	0.1



**Supplementary Figure 1.** The flowchart detailed the procedure for main hyperparameters selection of XGBoost models. We used “GridSearch” GridSearch s.GBoost models.XGBoost models.dhyperparameters (including gamma, alpha, max depth, nrounds, eta, min\_child\_weight, etc.). The possible values for each parameter are ranked and combined, and all possible combinations are listed to generate a “grid”. For each hyperparameter combination, train the model and use cross-validation to take the best performing hyperparameter combination as the best choice.

## Machine learning-based prognostic models of lung cancer brain metastases

**Supplementary Table 2.** Univariate and multivariate Cox analysis of LCBM characteristics extracted from training data

	OS					
	Univariate Cox analysis			Multivariate Cox analysis		
	HR	95% CI	P Value	HR	95% CI	P Value
<b>Age</b>						
< 50	Reference			Reference		
50-59	1.33	1.20-1.47	***	1.21	1.08-1.35	***
60-69	1.50	1.36-1.66	***	1.33	1.19-1.47	***
70-79	1.88	1.69-2.08	***	1.61	1.44-1.80	***
≥ 80	2.41	2.14-2.73	***	1.68	1.47-1.93	***
<b>Sex</b>						
Female	Reference			Reference		
Male	1.29	1.23-1.34	***	1.23	1.17-1.30	***
<b>Race</b>						
White	Reference			Reference		
Black	1.07	0.99-1.14	0.07	1.00	0.93-1.09	0.89
Others	0.68	0.63-0.74	***	0.71	0.65-0.77	***
<b>Marriage status</b>						
Married	Reference			Reference		
Divorced/Separated	1.28	1.20-1.37	***	1.24	1.15-1.33	***
Single	1.16	1.09-1.23	***	1.07	1.00-1.14	0.07
Widowed	1.33	1.24-1.43	***	1.12	1.03-1.22	**
<b>Months from diagnosis to therapy</b>						
0 month	Reference			Reference		
≥ 1 month	0.82	0.78-0.86	***	0.78	0.74-0.82	***
<b>Pathological type</b>						
Adenocarcinoma	Reference			Reference		
Squamous cell	1.71	1.60-1.83	***	1.48	1.37-1.59	***
Small cell	1.37	1.27-1.48	***	1.31	1.18-1.46	***
Large cell	1.29	1.11-1.51	**	1.38	1.17-1.64	***
Others	1.50	1.41-1.59	***	1.30	1.21-1.40	***
<b>Grade</b>						
Grade I	Reference			Reference		
Grade II	1.06	0.92-1.21	0.42	1.14	0.98-1.31	0.08
Grade III	1.42	1.24-1.61	***	1.36	1.18-1.56	***
Grade IV	1.52	1.31-1.76	***	1.39	1.17-1.65	***
<b>T Stage</b>						
T1	Reference			Reference		
T2	1.31	1.20-1.42	***	1.34	1.23-1.47	***
T3	1.53	1.41-1.67	***	1.50	1.37-1.65	***
T4	1.57	1.45-1.71	***	1.50	1.37-1.65	***
<b>N Stage</b>						
N0	Reference			Reference		
N1	1.06	0.97-1.16	0.18	1.05	0.95-1.15	0.36
N2	1.19	1.12-1.26	***	1.15	1.08-1.23	***
N3	1.23	1.14-1.31	***	1.18	1.09-1.28	***
<b>Chemotherapy</b>						
No/unknown	Reference			Reference		
Yes	0.36	0.34-0.38	***	0.39	0.37-0.41	***

## Machine learning-based prognostic models of lung cancer brain metastases

Radiotherapy						
No/unknown	Reference			Reference		
Yes	0.59	0.56-0.62	***	0.92	0.86-0.99	*
Surgery						
No	Reference			Reference		
Yes	0.48	0.43-0.53	***	0.54	0.48-0.60	***
Bone metastasis						
No	Reference			Reference		
Yes	1.14	1.09-1.20	***	1.20	1.13-1.27	***
Liver metastasis						
No	Reference			Reference		
Yes	1.46	1.37-1.55	***	1.34	1.25-1.44	***
Lung metastasis						
No	Reference			Reference		
Yes	1.24	1.17-1.3	***	1.13	1.06-1.20	***

\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

## Machine learning-based prognostic models of lung cancer brain metastases

**Supplementary Table 3.** Relevant characteristics of LCBM patients included from our hospital for XGBoost models external validation

Age	Sex	Marital	Race	Months	Pathological	Grade	T	N	Surgery	Radiation	Chemotherapy	Bone	Liver	Lung	6-month status	1-year status	2-year status
53	1	1	3	1	1	NA	NA	3	0	0	1	0	0	0	1	0	0
62	1	1	3	1	NA	NA	NA	3	0	0	1	1	1	1	0	0	0
56	2	1	3	1	3	NA	NA	3	1	0	1	0	0	0	1	1	0
69	2	1	3	1	1	NA	1	3	0	0	1	1	0	0	1	0	0
60	2	1	3	1	1	2	NA	NA	1	1	1	0	0	0	1	1	1
44	2	1	3	1	3	NA	NA	NA	0	1	0	0	0	0	0	0	0
63	2	1	3	1	2	NA	1	3	0	1	1	0	0	0	1	0	0
58	2	1	3	1	1	NA	NA	NA	0	1	1	0	0	0	1	0	0
61	2	1	3	1	NA	NA	NA	3	0	0	0	1	0	1	0	0	0
56	2	1	3	1	1	NA	NA	1	1	0	1	0	0	0	1	1	0
61	2	1	3	1	3	NA	NA	3	0	0	1	0	1	0	1	0	0
47	2	1	3	1	NA	NA	NA	NA	0	0	1	0	1	1	0	0	0
44	2	1	3	1	2	NA	4	3	0	0	0	0	0	1	0	0	0
45	1	1	3	1	3	NA	NA	3	0	1	1	0	0	0	1	0	0
72	2	1	3	1	NA	NA	3	NA	1	1	1	0	0	0	1	0	0
59	1	1	3	1	1	NA	4	NA	0	0	1	0	0	0	0	0	0
66	2	1	3	1	1	NA	NA	3	0	0	0	0	0	0	0	0	0
42	1	1	3	1	1	NA	1	3	0	0	1	0	0	0	0	0	0
75	1	4	3	1	1	3	NA	NA	0	1	0	0	0	0	0	0	0
62	2	1	3	1	2	NA	2	NA	0	1	1	1	0	1	0	0	0
54	2	1	3	1	5	1	NA	3	0	0	1	0	0	1	1	0	0
60	2	1	3	1	1	NA	4	0	0	0	1	1	1	1	0	0	0
55	1	1	3	1	1	NA	NA	1	1	1	1	1	0	0	1	1	1
69	2	1	3	1	5	NA	NA	NA	1	0	1	1	0	0	1	0	0
61	2	1	3	1	1	NA	NA	0	1	1	1	1	0	0	1	1	1
54	2	1	3	1	5	3	NA	0	1	0	0	0	0	0	0	0	0
72	2	1	3	1	2	3	NA	NA	0	0	0	0	0	0	0	0	0
56	2	1	3	2	2	NA	NA	NA	0	0	1	0	0	0	1	1	0
73	2	1	3	1	NA	NA	NA	NA	1	1	1	1	0	0	1	0	0
64	2	1	3	1	3	NA	NA	NA	1	0	1	0	0	0	1	0	0
52	2	1	3	1	3	NA	NA	NA	0	0	1	0	0	0	1	1	0
49	2	1	3	1	3	NA	NA	NA	0	0	1	0	0	0	1	1	0