

Supplementary Information

Supplementary Table 1 Demographic and clinicopathological characteristics of participants in the training and validation cohorts

	Total (N=1087)	Training cohort (N=761)	Validation cohort (N=326)	P value ^a
Age at diagnosis, median (IQR, years)	45 (25)	46 (24)	44 (25)	0.358
BMI, median (IQR, kg/m²)	22.4 (4.0)	22.4 (4.1)	22.4 (4.0)	0.845
Menstrual status, N (%)				
Premenopausal	681 (62.6%)	471 (61.9%)	210 (64.4%)	0.430
Postmenopausal	406 (37.4%)	290 (38.1%)	116 (35.6%)	
Pathological diagnosis, N (%)				
Benign	703 (64.7%)	494 (64.9%)	209 (64.1%)	0.799
Malignant	384 (35.3%)	267 (35.1%)	117 (35.9%)	
Histopathology of Benign masses, N (%)				
Endometriotic cysts	226 (32.1%)	162 (32.8%)	64 (30.6%)	0.815
Mature teratoma	185 (26.3%)	125 (25.3%)	60 (28.7%)	
Mucinous cystadenoma	95 (13.5%)	71 (14.4%)	24 (11.5%)	
Simple cyst	79 (11.2%)	56 (11.3%)	23 (11.0%)	
Serous cystadenoma	77 (11.0%)	53 (10.7%)	24 (11.5%)	
Others ^b	41 (5.8%)	27 (5.5%)	14 (6.7%)	
Histopathology of malignant masses, N (%)				
Serous	280 (72.9%)	195 (73.0%)	85 (72.6%)	0.871
Endometrioid	30 (7.8%)	21 (7.9%)	9 (7.7%)	
Mucinous	24 (6.3%)	15 (5.6%)	9 (7.7%)	
Others ^c	50 (13.0%)	36 (13.5%)	14 (12.0%)	
FIGO stage, N (%)				
I/ II	156 (40.6%)	104 (39.0%)	52 (44.4%)	0.313
III/ IV	228 (59.4%)	163 (61.0%)	65 (55.6%)	
Differential grade, N (%)				
G1/ G2	124 (32.3%)	93 (34.8%)	31 (26.5%)	0.114
G3	236 (61.5%)	158 (59.2%)	78 (66.7%)	
Unknown	24 (6.2%)	16 (6.0%)	8 (6.8%)	

Continuous variables were presented as median (IQR) and analyzed using the Mann–Whitney U test. Categorical variables were presented as N (%) and analyzed using the Pearson Chi-Square test. Abbreviations: IQR, interquartile range; BMI, body mass index; FIGO, International Federation of Gynecology and Obstetrics.

^a P value for comparison between the training and validation cohorts.

^b Including fibroma, thecoma, Brenner tumor, benign struma ovarii.

^c Including clear cell carcinoma, malignant Brenner tumor, carcinosarcoma, endometrioid stromal sarcoma, adult granulosa cell tumor, immature teratoma, malignant Sertoli-Leydig cell tumors, malignant struma ovarii, mucinous carcinoid, small round cell malignant tumor.

Supplementary Table 2 Laboratory and ultrasound characteristics of participants in the training and validation cohorts

	Total (N=1087)	Training cohort (N=761)	Validation cohort (N=326)	P value ^a
Laboratory characteristics				
CA125 (U/ml)	32.5 (153.8)	30.8 (140.2)	35.0 (187.7)	0.426
HE4 (pmol/L)	52.8 (59.6)	53.0 (56.4)	51.7 (62.5)	0.888
HGB (g/L)	135 (21)	135 (21)	136 (20)	0.969
HCT (%)	41.1 (5.4)	41.1 (5.5)	41.0 (5.1)	0.920
MCV (fL)	90.7 (7.4)	90.8 (7.6)	90.7 (7.0)	0.777
Neutrophils (10 ⁹ /L)	3.52 (1.96)	3.51 (2.05)	3.57 (1.82)	0.742
Monocytes (10 ⁹ /L)	0.32 (0.17)	0.32 (0.17)	0.32 (0.16)	0.604
Lymphocytes (10 ⁹ /L)	1.48 (0.64)	1.48 (0.64)	1.49 (0.69)	0.269
Basophils (10 ⁹ /L)	0.03 (0.01)	0.02 (0.01)	0.02 (0.01)	0.727
PLT (10 ⁹ /L)	243 (103)	243 (101)	242 (110)	0.705
MPV (fL)	10.8 (1.9)	10.8 (2.0)	10.9 (1.8)	0.469
ALB (g/L)	45.0 (4.9)	45.1 (4.9)	44.9 (5.0)	0.770
FIB (g/L)	2.88 (1.01)	2.86 (0.97)	2.93 (1.08)	0.538

NLR	2.28 (1.80)	2.28 (1.81)	2.26 (1.78)	0.675
PLR	160 (97)	160 (101)	159 (94)	0.619
MLR	0.21 (0.14)	0.21 (0.14)	0.21 (0.15)	0.336
FAR	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.581
SII	548 (558)	544 (553)	561 (573)	0.791
Ultrasound characteristics				
Maximum diameter (mm)	73.0 (51.0)	74.0 (54.0)	73.0 (44.0)	0.774
Bilateral masses				
No	842 (77.5%)	590 (77.5%)	252 (77.3%)	0.934
Yes	245 (22.5%)	171 (22.5%)	74 (22.7%)	
Irregular masses				
No	809 (74.4%)	564 (74.1%)	245 (75.2%)	0.719
Yes	278 (25.6%)	197 (25.9%)	81 (24.8%)	
Presence of solid areas				
No	530 (48.8%)	371 (48.8%)	159 (48.8%)	0.995
Yes	557 (51.2%)	390 (51.2%)	167 (51.2%)	
Strong blood flow signal				
No	836 (76.9%)	582 (76.5%)	254 (77.9%)	0.607
Yes	251 (23.1%)	179 (23.5%)	72 (22.1%)	

Continuous variables were presented as median (IQR) and analyzed using the Mann–Whitney U test. Categorical variables were presented as N (%) and analyzed using the Pearson Chi-Square test. Abbreviations: CA125, cancer antigen 125; HE4, human epididymis protein 4; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; PLT, platelet; MPV, mean platelet volume; ALB, albumin; FIB, fibrinogen; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; FAR, fibrinogen-to-albumin ratio; SII, systemic immune-inflammation index.

^a P value for comparison between the training and validation cohorts.

Supplementary Table 3 Values of the penalty parameter lambda (λ) in the LASSO regression and the corresponding coefficients of candidate variables

	lambda.min (0.002424828)	lambda.1se (0.02481887)
Demographic characters		
Age at diagnosis (years)	0.044	0.022
BMI (kg/m ²)	0.073	0
Tumor markers		
CA125 > cutoff (100 U/ml)	2.186	1.249
HE4 > cutoff (64.5 pmol/L)	1.822	1.323
Systemic inflammatory response markers		
HGB > cutoff (133 g/L)	0	0
MCV > cutoff (81.9 fl)	0.355	0
HCT > cutoff (38.7%)	-0.317	0
Neutrophils > cutoff (3.82*10 ⁹ /L)	1.054	0
Monocytes > cutoff (0.40*10 ⁹ /L)	-0.719	0
Lymphocytes > cutoff (1.37*10 ⁹ /L)	-0.325	0
Basophils > cutoff (0.02*10 ⁹ /L)	-0.171	0
PLT > cutoff (285*10 ⁹ /L)	0.701	0
MPV > cutoff (10.7 fl)	-0.008	0
ALB > cutoff (42.3 g/L)	-0.235	0
FIB > cutoff (3.28 g/L)	1.473	0.948
NLR > cutoff (2.80)	-0.477	0
PLR > cutoff (200)	0.291	0
MLR > cutoff (0.25)	2.215	1.131
FAR > cutoff (0.07)	-0.443	0
SII > cutoff (698)	-0.923	0
Ultrasound characteristics		
Maximum diameter (mm)	0.002	0
Bilateral masses	-1.352	0

Irregular masses	0.928	0.467
Presence of solid areas	1.107	0.687
Strong blood flow signal	2.463	1.402
Intercept	-7.926	-4.207

Abbreviations: LASSO, least absolute shrinkage and selection operator; BMI, body mass index; CA125, cancer antigen 125; HE4, human epididymis protein 4; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; PLT, platelet; MPV, mean platelet volume; ALB, albumin; FIB, fibrinogen; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; FAR, fibrinogen-to-albumin ratio; SII, systemic immune-inflammation index.

Supplementary Table 4 Multivariate logistic regression analysis with all the variables selected from LASSO regression

Variables	β	SE	Wald test	OR (95% CI)	P value
Age at diagnosis, years	0.04	0.01	10.75	1.04 (1.02, 1.07)	0.001
CA125 > cutoff (100 U/ml)	2.12	0.40	27.50	8.34 (3.83, 18.82)	< 0.001
HE4 > cutoff (64 pmol/L)	1.72	0.38	20.37	5.58 (2.66, 11.89)	< 0.001
FIB > cutoff (3.28 g/L)	1.24	0.33	14.15	3.46 (1.81, 6.62)	< 0.001
MLR > cutoff (0.25)	2.02	0.34	35.36	7.51 (3.92, 14.90)	< 0.001
Irregular masses	0.72	0.41	3.10	2.06 (0.91, 4.59)	0.078
Presence of solid areas	1.42	0.33	18.32	4.32 (2.29, 8.40)	< 0.001
Strong blood flow signal	2.39	0.43	30.24	10.92 (4.73, 26.16)	< 0.001

Abbreviations: β , coefficient; SE, standard error; OR, odds ratio; CI, confidence interval; CA125, cancer antigen 125; HE4, human epididymis protein 4; FIB, fibrinogen; MLR, monocyte-to-lymphocyte ratio.

Supplementary Table 5 The Δ AUC, NRI, and IDI of three built models for discriminating early-stage ovarian cancer from benign ovarian masses

	Model 2 – Model 1	Nomogram – Model 1	Nomogram – Model 2
ΔAUC	0.023	0.052	0.029
P value	0.01	< 0.001	< 0.001
NRI	0.05 ^a	0.13 ^b	0.11 ^b
95% CI	-0.02-0.12	0.08-0.18	0.06-0.16
P value	0.158	< 0.001	< 0.001
IDI	0.08	0.16	0.08
95% CI	0.05-0.11	0.12-0.20	0.05-0.10
P value	< 0.001	< 0.001	< 0.001

Abbreviations: AUC, area under the curve; NRI, net reclassification improvement; IDI, integrated discrimination improvement; CI, confidence interval.

^a cutoff = 0.393, ^b cutoff = 0.180.

Supplementary Table 6 The coding of variables

Variables	Coding
Diagnosis	
Benign	0
Malignant	1
Pathology of Benign ovarian masses	
Endometrioma	1
Mature teratoma	2
Mucinous cystadenoma	3
Simple cyst	4
Serous cystadenoma	5
Other Benign ovarian masses	6
Pathology of malignant ovarian masses	
Serous	7
Endometrioid	8
Mucinous	9

Other malignant ovarian masses	10
Differential grade	
G1/G2	1
G3	2
FIGO stage	
I/ II	1
III/ IV	2
Menstrual status	
Premenopausal	0
Postmenopausal	1
Bilateral masses	
No	0
Yes	1
Irregular masses	
No	0
Yes	1
Presence of solid areas	
No	0
Yes	1
Strong blood flow signal	
No	0
Yes	1

Abbreviations: BMI, body mass index; FIGO, International Federation of Gynecology and Obstetrics.

Supplementary Table 7 R packages used in this research

R packages	Version	Application
glmnet	4.1-3	LASSO regression
forestplot	2.0.1	Drawing forest plot
pROC	1.18.0	ROC analysis
nrccens	1.6	Calculating NRI and getting NRI figures
PredictABEL	1.2-4	Calculating NRI and IDI; getting 95% CI and P values of NRI/IDI
ResourceSelection	0.3-5	Hosmer-Lemeshow test
rmca	1.6	Drawing DCA curve
rms	6.2-0	Drawing nomogram
regplot	1.1	Drawing dynamic nomogram
DynNom	5.0.1	Generating online dynamic nomogram

Abbreviations: LASSO, least absolute shrinkage and selection operator; OR, odds ratio; ROC, receiver operating characteristic; NRI, net reclassification improvement; IDI, integrated discrimination improvement; CI, confidence interval; DCA, decision curve analysis.