RESEARCH ARTICLE

A novel prognostic score based on carbohydrate antigen 125, alpha-fetoprotein, and carcinoembryonic antigen for predicting postoperative prognosis in endometrial cancer: results from a retrospective cohort study

Zi-hao Wang[1†], Qing Li[1†], Jiahui Gu[1†], Yun-zheng Zhang[1], Shu-wen Ge[1], Lu-he Shan[1], Bo Wang[1], Zi-yu Zhang[1], Qi-jun Wu[2], Hui Yang[1*], Xiao-xin Ma[1*]

[1]Department of Obstetrics and Gynecology, Key Laboratory of Obstetrics and Gynecology of Higher Education of Liaoning Province, Shengjing Hospital of China Medical University, Shenyang, China; [2]Department of Clinical Epidemiology, Shengjing Hospital of China Medical University, Shenyang, China

Abstract

Background: Endometrial cancer (EC) is a common gynecological malignancy with increasing incidence. While several serum biomarkers have been studied for EC, their combined prognostic value remains unclear. This study aimed to evaluate the prognostic significance of preoperative serum CA-125, CA-19-9, CA-72-4, CEA, and AFP levels in EC patients and develop a risk score for predicting survival outcomes.

Methods: A retrospective cohort study of2081 EC patients was conducted at Shengjing Hospital of China Medical University. Serum biomarker levels and clinicopathological data were collected. Univariate and multivariate Cox proportional hazard models were used to identify independent prognostic factors. A risk score based on CA125, CEA, and AFP levels was developed using LASSO-Cox regression, and nomograms were constructed for survival prediction.

Results: Multivariate analysis identified elevated CA125 ($p=0.003$), AFP ($p<0.0001$), and CEA ($p=0.014$) as independent factors for overall survival (OS). These markers were also independent predictors of progression-free survival (PFS). The risk score incorporating CA125, AFP, and CEA was an independent indicator for both PFS ($p<0.0001$) and OS ($p<0.0001$). Nomograms based on the risk score and clinicopathological features demonstrated good predictive ability and calibration for survival outcomes.

Conclusions: The risk score based on preoperative serum levels of CA125, CEA, and AFP is a valuable prognostic tool for predicting PFS and OS in EC patients. Nomograms incorporating this risk score accurately predict EC prognosis and may aid in clinical decision-making.

Funding: This study was supported by the National Key R&D Program of China (Program Nos. 2022YFC2704400), The National Natural Science Foundation of China (No. 81872123 and 81472438); University innovation team of Liaoning Province; Special Professor of Liaoning Province 'Major Special Construction Plan' for Discipline Construction of China Medical University in 2018 (No. 3110118029); Outstanding Scientific Fund of Shengjing Hospital (No. 201601) 'Major Special Construction Plan' for Discipline Construction of China Medical University in 2018.

***For correspondence:** yangh9@sj-hospital.org (HY); maxx@sj-hospital.org (X-xM)

[†]These authors contributed equally to this work

Competing interest: The authors declare that no competing interests exist.

Funding: See page 19

Received: 22 November 2023
Preprinted: 26 January 2024
Accepted: 20 March 2026
Published: 10 August 2026

Reviewing Editor: Yongliang Yang, Shanghai University of Medicine and Health Sciences, China

© Copyright Wang, Li, Gu *et al*. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Editor's evaluation

This study presents a valuable finding on a potential efficient model for the prognosis prediction of patients with endometrial cancer (EC). The evidence supporting the claims of the authors is solid. The work will be of interest to scientists working in the field of endometrial cancer (EC).

Introduction

Endometrial cancer (EC) is one of the most common gynecological cancers in developed countries such as Europe and the United States, accounting for approximately 20–30% of all gynecological tumors. Recent years have also seen year-on-year increases in tumor incidence in Asian countries (***Crosbie et al., 2022***; ***Miller and King, 2008***). In 2021, there were 66,570 new cases of EC in the United States and 12,940 deaths related to it (***Siegel et al., 2021***). With increases in life expectancy and rates of obesity, corresponding increases in the frequency of EC worldwide are expected (***Angioli et al., 2013***). EC ranks second among the gynecological malignancies in China (***Chen et al., 2016***). An early symptom of EC is postmenopausal vaginal bleeding; most patients can be diagnosed early. However, there are still more than 28% of patients diagnosed at an advanced stage (***Moore et al., 2011***). And 10–20% of early-stage EC patients have recurrence or metastasis after treatment, which has become the main cause of their death (***Takahashi et al., 2017***). Currently, potential factors to determine prognosis (tumor grade, staging) are being closely monitored in an effort to predict EC patient survival. However, some patients presenting with similar clinicopathological features experience very different outcomes (***Makker et al., 2021***). As such, we sought to identify effective EC biomarkers that are able to predict prognosis in EC patients and identify potential patients at high risk of developing EC.

A tumor marker refers to a substance produced and secreted by tumor cells that subsequently enter body fluids or tissues (***Virji et al., 1988***). At present, the mechanisms of tumorigenesis and development are not completely clear, and identification of tumor markers is paramount. With recent developments in the field of tumor biology and immunology, some serum tumor markers that may potentially make a significant impact on the initiation and progression of gynecological malignancies have been discovered (***Dochez et al., 2019***; ***Felder et al., 2014***; ***Laengsri et al., 2018***). High-risk populations and women with systemic and advanced disease may benefit from early detection using serum markers. Sensitive and specific tumor serum markers would also provide a means of monitoring treatment response and enable prediction of recurrence or metastasis.

Five serological tumor markers, including carbohydrate antigen 125 (CA125), carbohydrate antigen 19–9 (CA19-9), carbohydrate antigen 72–4 (CA72-4), carcinoembryonic antigen (CEA), and alpha-fetoprotein (AFP), are routinely used in clinical practice as tools for cancer diagnosis and prediction of prognosis. The most widely used biomarker for EC is CA125. It is used to early diagnose, detect tumor recurrence, and to predict survival outcomes (***Niloff et al., 1984***; ***Dotters, 2000***; ***Jiang et al., 2015***; ***Nicklin et al., 2012***). However, the ability of CA125 to specifically reflect the metastatic potential of tumors is still insufficient. CA125 levels will not increase during the progression of some patients. Moreover, numerous EC serological tumor markers have been studied, including HE4 (human epididymal protein 4), CA19-9, AFP, and CEA (***Cheng et al., 2021***; ***Kozakiewicz et al., 2016***). Serum HE4 can be used as an independent prognostic factor for patients with EC (***Behrouzi et al., 2021***). Moreover, the predictive efficacy of HE4 for disease recurrence seems to be superior to CA125 (***Espiau Romera et al., 2021***), especially in high-risk EC patients with high incidence of distant metastasis (***Abbink et al., 2018***). Bian et al. showed that combined use of HE4, CA125, CA72-4, and CA19-9 is valuable in the early diagnosis of EC and that these markers are useful immunological tissue markers for patients with EC (***Bian et al., 2017***). Unfortunately, serum markers are more commonly used for evaluation of diagnosis for EC patients. Identification of other specific serum cancer biomarkers has important clinical significance for the early diagnosis of recurrence or metastasis. The study of EC tumor markers and their expression pattern can provide additional information when making judgments about EC pathology and course of treatment. It plays a significant role in patient evaluation and progress of individualized treatment; it is also especially crucial for prediction of prognosis. We performed a retrospective study at Shengjing Hospital of China Medical University. The aim was to assess the relationship between pre-operative serological tumor biomarkers and clinicopathological

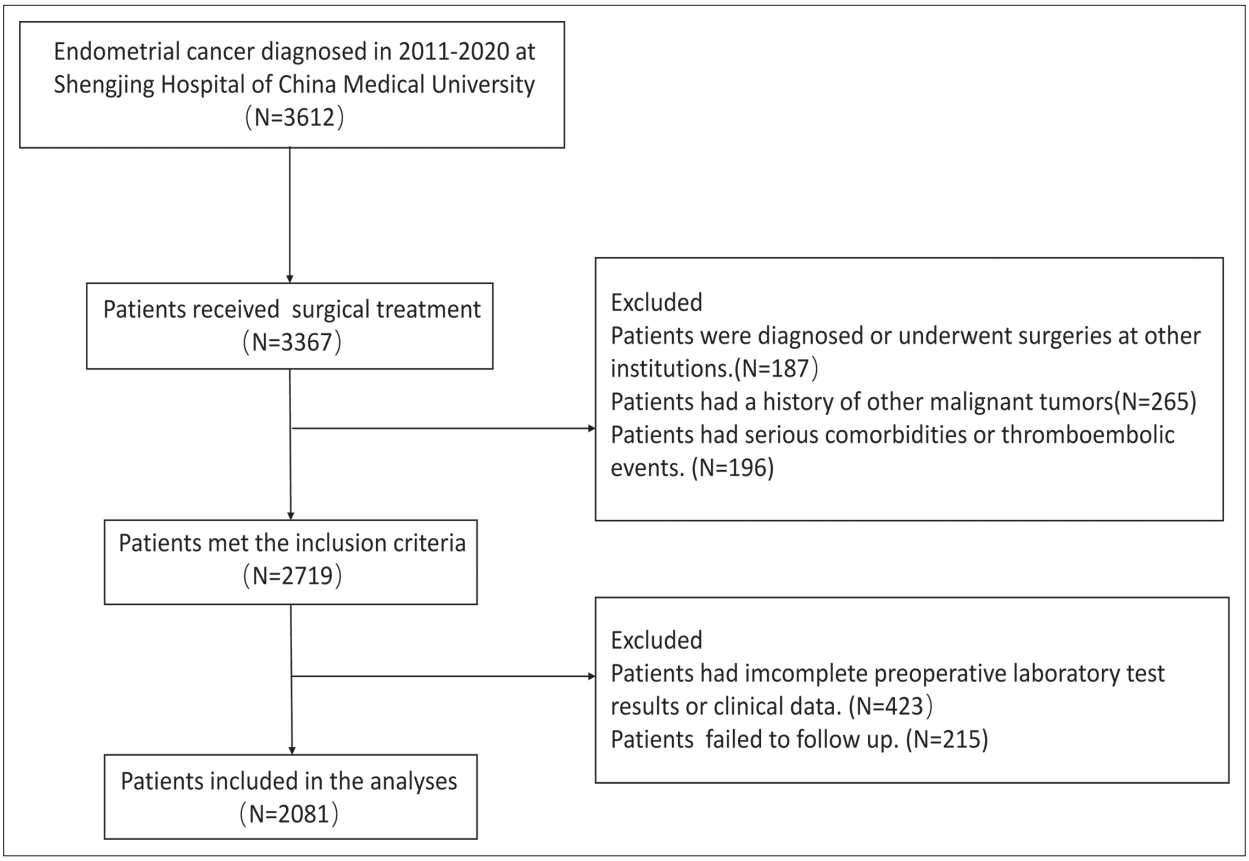

Figure 1. Flow chart of the patient inclusion and exclusion process.

parameters in EC patients and determine novel biomarkers that are associated with post-operative survival.

However, the predictive ability of single-serum tumor markers remains insufficient. We aim to explore a new signature involving multiple serum tumor markers that may improve the prognosis prediction.

Materials and methods

Study population

We performed a retrospective study of EC patients from December 2010 to March 2020 at Shengjing Hospital affiliated to China Medical University. This study was approved by the Institutional Review Committee of the Shengjing Hospital of China Medical University (No. 2017PS292K). All patients included in the study met the following criteria: (1) age ≥18 years, (2) underwent hysterectomy at Shengjing Hospital, (3) post-operative diagnosis of EC by pathology, (4) no history of other malignant tumors, (5) no active infection, (6) no blood disease or other serious comorbidities, (7) no thrombo-embolic events, and (8) complete data for all required variables in this study. A total of 1286 patients did not meet all the inclusion criteria and were excluded, and 2081 patients were included (***Figure 1***). The study complied with the Declaration of Helsinki and was approved by the Ethics Committee of Shengjing Hospital of China Medical University (No. 2017PS292K). Informed consent was acquired from all participants of research.

Sample-size calculation

The sample size was calculated using an online sample calculation website, http://riskcalc.org:3838/samplesize/. According to previous research, the abnormality of AFP level had the lowest prevalence and had a prevalence of approximately 1.8%. We expected that patients with high AFP levels would

have an HR value of 3, compared to those with normal AFP levels. The Type I error rate set in this study was 0.05, the degree of certainty of the study was 0.8, and the approximate ratio of the sample size between the three populations was 1. A minimum sample size of 1680 was calculated for this study.

Data collection

The following demographic and clinical data were collected from the electronic medical records of the Shengjing Hospital Information System: patient's age at diagnosis, pre/post-menopausal status, International Federation of Obstetrics and Gynecology Association (FIGO) stage, tumor grade, histology type, tumor myometrial invasion depth, and existing comorbidities. Pre-operative levels of serum tumor markers (including CA125, CA72-4, CA19-9, AFP, and CEA) in all enrolled patients were measured using a chemiluminescence immunoassay.

Data collection was conducted by professional gynecologists and pathologists. Evaluation and classification of tumors was in accordance with the WHO classification during post-operative examination of tumor pathology, and all tumors were clinically staged according to FIGO guidelines (2009). Pathologists divide tumors into well (G1), moderately (G2), or poorly (G3) differentiated, and tumor histology was classified as type I (mainly endometrioid) or type II (non-endometrioid).

Follow-up and outcome

We conduct follow-up of all patients uniformly in January 2023. Patient follow-up information was obtained from medical records or telephone interviews. Progression-free survival (PFS) was defined as the time from surgery to EC progression or death, and overall survival (OS) was defined as the time between surgery and death. The date of death was obtained during follow-up or from a death certificate provided.

Statistical analysis

The data input was checked to ensure its accuracy. All data were statistically analyzed using SPSS ver. 20.0 (IBM Corp, Armonk, NY). Descriptive statistics were used to represent patients' general characteristics. The chi-square test and the t-test were respectively used for qualitative and quantitative variables to identify associations between clinicopathological features and tumor biomarkers. A Kaplan–Meier analysis was used to compare differences in survival, and a log-rank test was used to make comparisons between sub-groups. Univariate and multivariate analyses were used to identify factors that related to patient survival, and the Cox proportional hazards model was used to test for independence of the effect. Data were expressed as hazard ratios (HRs) and 95% confidence intervals (CIs). Further, the Lasso (least absolute shrinkage and selection operator) regression analysis was performed to obtain a risk score formula based on serum expression levels of the three tumor markers. Based on the clinical prediction model, the prognostic nomogram was constructed using R version 4.2.0. The C index was used to assess the predictive power of the model, and the calibration curve was used to evaluate the calibration degree of the clinical prediction model.

Results

Patient baseline characteristics

Characteristics of eligible patients are summarized in ***Table 1***. Of the 2081 patients, 1964 had tumors classified as pathological type I, and 117 had tumors classified as type II. 600 patients had not experienced menopause at the time of diagnosis, and the remaining 1481 cases were diagnosed with EC after menopause. Post-operative FIGO staging by examination of EC pathology identified 1645 stage I cases and 436 stage II–IV cases. 1052 patients had highly differentiated tumors, and 1029 patients had moderately and poorly differentiated tumors. The seroprevalence rates of pre-operative CA125, CA19-9, CA72-4, AFP, and CEA were 27.3% (569/2081), 17.2% (357/2081), 25.7% (534/2081), 1.73% (36/2081), and 5.09% (106/2081) respectively.

Serum CA125, CA19-9, CA72-4, CEA, and AFP levels for patients with EC

Classified patients into two groups, event (recurrence/metastasis/death occurred) and non-event (none of the above occurred) groups. The two groups were compared for levels of pre-operatively

Table 1. Patient characteristics and relationships between the clinicopathological characteristics and serum CA125, CA19-9, CA72-4, AFP, and CEA.

	Cases	CA125			CA199			CA724			AFP			CEA		
	No. (%)	≤ 35	>35	p	≤ 37	>37	p	≤ 6.9	>6.9	p	≤ 9	>9	p	≤ 5	>5	p
	2081	1512	569		1724	357		1547	534		2045	36		1975	106	
Age				0.933			0.405			<0.001			0.029			0.0003
<60	1288 (61.9)	935 (61.8)	353 (62.04)		1074 (62.3)	214 (59.94)		1010 (65.3)	278 (52.1)		1272 (62.2)	16 (44.4)		1237 (62.6)	51 (48.1)	
≥60	793 (38.1)	577 (38.2)	216 (37.96)		650 (37.7)	143 (40.06)		537 (34.7)	256 (47.9)		773 (37.8)	20 (55.6)		738 (37.4)	55 (51.9)	
BMI				0.089			0.022			0.027			0.52			0.116
<28	1707 (82)	1227 (81)	480 (84.4)		1399 (81.1)	308 (86.3)		1252 (80.9)	455 (85.2)		1676 (82)	31 (86.1)		1614 (81.7)	93 (87.7)	
≥28	374 (18)	285 (19)	89 (15.6)		325 (18.9)	49 (13.7)		295 (19.1)	79 (14.8)		369 (18)	5 (13.9)		361 (18.3)	13 (12.3)	
Menopause				0.051			0.003			0.582			0.818			<0.001
No	600 (28.8)	418 (27.6)	182 (32)		520 (30.2)	80 (22.3)		451 (29.2)	149 (27.9)		589 (28.8)	11 (30.6)		585 (29.6)	15 (14.2)	
Yes	1481 (71.2)	1094 (72.4)	387 (68)		1204 (69.8)	277 (77.8)		1096 (70.8)	385 (72.1)		1456 (71.2)	25 (69.4)		1390 (70.4)	91 (85.8)	
Histological type				0.019			0.601			0.01			0.004			0.08
I	1964 (94.4)	1438 (95.1)	526 (92.4)		1625 (94.3)	339 (95)		1472 (95.2)	492 (92.1)		1934 (94.6)	30 (83.3)		1868 (94.6)	96 (90.6)	
II	117 (5.6)	74 (4.9)	43 (7.6)		99 (5.7)	18 (5)		75 (4.8)	42 (7.9)		111 (5.4)	6 (16.7)		107 (5.4)	10 (9.4)	
Differentiation degree				0.149			0.956			0.425			0.687			0.012
G1	1052 (50.6)	779 (51.5)	273 (48)		872 (50.6)	180 (50.4)		790 (51.1)	262 (49.1)		1035 (50.6)	17 (47.2)		1011 (51.2)	41 (38.7)	
G2-G3	1029 (49.4)	733 (248.5)	296 (52)		852 (49.4)	177 (49.6)		757 (48.9)	272 (50.9)		1010 (49.4)	19 (52.8)		964 (48.8)	65 (61.3)	
Myometrial invasion				<0.001			<0.001			0.007			0.502			<0.001
Myometrial <1/2	1654 (79.5)	1256 (83.1)	398 (69.9)		1397 (81.0)	257 (72.0)		1251 (80.9)	403 (75.5)		1627 (79.6)	27 (75)		1595 (80.8)	59 (55.7)	
Myometrial ≥1/2	427 (20.5)	256 (16.9)	171 (30.1)		327 (19.0)	100 (18.0)		296 (19.1)	131 (24.5)		418 (20.4)	9 (25)		380 (19.2)	47 (44.3)	
FIGO stage				<0.001			<0.001			<0.001			<0.001			<0.001
I	1645 (79.1)	1297 (85.8)	348 (61.2)		1406 (81.6)	239 (66.9)		1259 (81.4)	386 (72.3)		1625 (79.5)	20 (55.6)		1585 (80.3)	60 (56.6)	
II–IV	436 (20.9)	215 (14.2)	221 (38.8)		318 (18.4)	118 (33.1)		288 (18.6)	148 (27.7)		420 (20.5)	16 (44.4)		390 (19.7)	46 (43.4)	

measured CA125, CA19-9, CA72-4, CEA, and AFP. Levels of CA125 ($p < 0.0001$), CA19-9 ($p=0.0005$), CA72-4 ($p=0.0021$), CEA ($p < 0.0001$), and AFP ($p < 0.0001$) were significantly higher in the event group compared to the non-event group (***Figure 2***). As such, CA125, CA19-9, CEA, and AFP may be used as potential markers to predict the prognosis of EC patients.

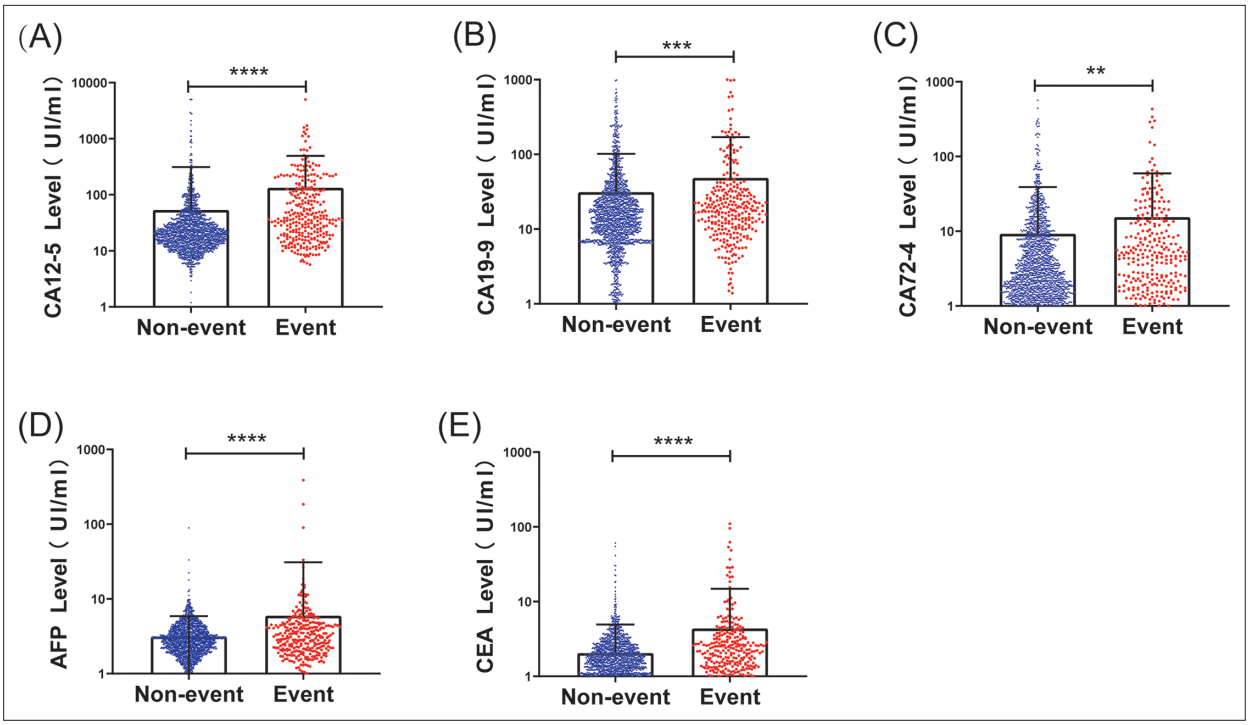

Figure 2. Comparison of preoperative serum biomarkers levels. Comparison of preoperative serum CA125 (**A**), CA19-9 (**B**), CA72-4 (**C**), AFP, (**D**) and CEA (**E**) levels in the event (recurrence/metastasis/death occurred) and non-event groups of endometrial cancer patients (N=906). * We defined p<0.05 as a statistical difference.

The online version of this article includes the following source data for figure 2:

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patient population in this study.

Univariate and multivariate survival analysis of EC patients

For this study, the median follow-up time was 48 months, 5-year OS was 90.0%, and 5-year PFS was 87.0%. Single factor survival analysis was performed to determine the effect of variables such as age, menopausal status, pathological type, degree of differentiation, FIGO stage, depth of myometrial invasion, hypertension, diabetes, complications, lymphovascular invasion (LVSI), expression of ER, PR, P53 in pathological tissues, serum CA125, CA19-9, CEA, and AFP levels in EC patients.

In univariate analysis (***Table 2***), age ≥60 (p<0.001), menopausal status (p<0.001), pathological type II (p<0.001), low degree of differentiation (p<0.0001), FIGO stage ≥II (p<0.001), depth of myometrial invasion ≥1/2 (p<0.001), positive lymphovascular infiltrates (p<0.001), negative ER expression (p<0.001), negative PR expression (p<0.001), and positive P53 expression (p=0.006), CA125 > 35 (p<0.001), CA72-4 > 6.9 (p≤0.001), CEA > 5 (p<0.001), and AFP > 9 (p≤0.001) were significantly associated with poor OS prognosis for EC patients. Serum levels of CA125 (p<0.001), CA72-4 (p≤0.001), CEA (p<0.001), and AFP (p≤0.001) were also associated with poor PFS prognosis for EC patients.

Differences in survival for serum CA125, CA19-9, CA72-4, CEA, and AFP were compared using a Kaplan–Meier analysis. We identified serum levels of CA125, CA72-4, AFP, and CEA as factors for prognosis of OS and PFS in EC patients (***Figure 3***).

A multivariate analysis of factors that had been identified as being important by univariate analysis was performed (***Table 2***). These factors included age, menopausal status, pathological type, degree of differentiation, FIGO stage, depth of myometrial invasion, lymphovascular invasion (LVSI), expression of ER, PR, P53 in pathological tissues, serum CA125, CA72-4, CEA, and AFP levels. Elevated serum CA125 (p<0.001) and AFP (p<0.001), CEA (p=0.037), age (p=0.023), menopausal status (p=0.034), pathological type (p=0.004), degree of differentiation (p=0.001), FIGO stage (p<0.001), and lymphovascular infiltrates (p<0.001) were independent factors for PFS prognosis in EC patients. Increased serum CA125 (p=0.003), AFP (p<0.001), CEA (p=0.014), age (p=0.004), pathological type (p=0.005), degree of differentiation (p=0.015), FIGO stage (p<0.001), and lymphovascular infiltrates (p=0.003) were independent factors for prognosis of OS in EC patients (***Table 2***).

Table 2. Overall survival and progression-free survival of the preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables.

	OS				PFS			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
Variables	p-Value	HR (95% CI)	p-Value	HR (95% CI)	p-Value	HR (95% CI)	p-Value	HR (95% CI)
AFP (U/mL)	<0.0001		<0.0001		<0.0001		<0.0001	
AFP <9		Ref		Ref		Ref		Ref
AFP ≥9		6.039 (3.737–9.758)		6.914 (3.374–14.165)		5.559 (3.490–8.855)		7.437 (3.717–14.879)
CEA (U/mL)	<0.0001		0.014		<0.0001		0.037	
CEA <5		Ref		Ref		Ref		Ref
CEA ≥5		3.21 (2.237–4.598)		2.34 (1.185–4.610)		3.549 (2.554–4.932)		1.967 (1.042–3.713)
CA125 (U/mL)	<0.0001		0.003		<0.0001		<0.0001	
CA125<35		Ref		Ref		Ref		Ref
CA125≥35		3.158 (2.480–4.020)		1.989 (1.270–3.116)		3.296 (2.634–4.124)		2.392 (1.593–3.590)
CA19-9 (U/mL)	0.108		0.162		0.152		0.062	
CA199<37		Ref		Ref		Ref		Ref
CA199≥37		1.28 (0.948–1.717)		0.665 (0.375–1.178)		1.226 (0.928–1.619)		0.611 (0.365–1.025)
CA72-4 (U/mL)	<0.0001		0.615		<0.0001		0.726	
CA724<6.9		Ref		Ref		Ref		Ref
CA724≥6.9		1.701 (1.315–2.201)		1.122 (0.717–1.757)		1.686 (1.329–2.140)		1.075 (0.716–1.614)
Age (years)	<0.0001		0.004		<0.0001		0.023	
<60		Ref		Ref		Ref		Ref
≥60		2.550 (1.996–3.257)		1.963 (1.233–3.127)		2.266 (1.808–2.839)		1.612 (1.067–2.435)
Fertility situation	0.082				0.625			
none								
yes		1.568 (0.945–2.602)				1.062 (0.834–1.354)		
Menopause	<0.0001		0.128		<0.0001		0.034	
no		Ref		Ref		Ref		Ref
yes		2.31 (1.649–3.242)		1.665 (0.863–3.212)		2.056 (1.530–2.765)		1.884 (1.048–3.386)
BMI (kg/m^2)	0.140				0.053			
<28		Ref				Ref		
≥28		1.244 (0.931–1.662)				1.301 (0.996–1.698)		
FIGO stage	<0.0001		<0.0001		<0.0001		<0.0001	
Ⅰ		Ref		Ref		Ref		Ref
Ⅱ-Ⅳ		6.214 (4.867–7.934)		3.199 (1.982–5.164)		6.854 (5.458–8.607)		3.888 (2.511–6.021)
Pathological type	<0.0001		0.005		<0.0001		0.004	
Type I		Ref		Ref		Ref		Ref
Type II		6.317 (4.628–8.621)		2.233 (1.282–3.888)		6.019 (4.502–8.048)		2.104 (1.268–3.493)
Myometrial invasion	<0.0001		0.494		<0.0001		0.437	
<1/2		Ref		Ref		Ref		Ref
≥1/2		2.982 (2.326–3.824)		1.183 (0.731–1.914)		3.003 (2.385–3.782)		1.187 (0.771–1.828)

Table 2 continued on next page

Table 2 continued

	OS				PFS			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
Variables	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**
Differentiation	<0.0001		0.015		<0.0001		0.001	
G1		Ref		Ref		Ref		Ref
G2/3		3.452 (2.602–4.582)		1.941 (1.137–3.315)		3.453 (2.658–4.485)		2.237 (1.363–3.672)
LVSI	<0.0001		0.003		<0.0001		0.002	
Negative		Ref		Ref		Ref		Ref
Positive		3.674 (2.388–5.652)		2.599 (1.390–4.860)		4.046 (2.752–5.949)		2.498 (1.419–4.397)
diabetes	0.169				0.068			
no		Ref				Ref		
yes		1.23 (0.916–1.653)				1.287 (0.981–1.687)		
hypertension	0.810				0.711			
no		Ref				Ref		
yes		1.031 (0.802–1.325)				0.957 (0.759–1.207)		
Comorbidities	0.191				0.183			
no		Ref				Ref		
yes		1.559 (0.802–3.030)				1.535 (0.817–2.882)		
expression of ER	<0.0001		0.991		<0.0001		0.953	
Negative		Ref		Ref		Ref		Ref
Positive		0.682 (0.556–0.836)		0.998 (0.680–1.463)		0.637 (0.530–0.767)		0.989 (0.694–1.411)
expression of PR	<0.0001		0.421		<0.0001		0.709	
Negative		Ref		Ref		Ref		Ref
Positive		0.569 (0.469–0.690)		0.851 (0.575–1.260)		0.559 (0.468–0.667)		0.934 (0.652–1.337)
expression of P53	0.006		0.823		<0.0001		0.452	
Negative		Ref		Ref		Ref		Ref
Positive		1.515 (1.129–2.033)		0.966 (0.715–1.305)		1.732 (1.323–2.267)		1.111 (0.845–1.461)

To assess the efficacy of selecting clinical test outliers for CA-125, CA19-9, CA72-4, AFP, and CEA, we employed ROC curves and the Youden index to determine cut-off values. Patients were grouped based on these cut-offs. Differences in survival for serum CA125, CA19-9, CA72-4, CEA, AFP were compared using a Kaplan–Meier analysis (***Figure 3—figure supplement 1***), and univariate and multivariate Cox regression analyses were performed to evaluate the prognostic impact of these biomarkers on EC. The results confirmed that the prognostic value of each marker remained consistent with prior analyses when using these cut-offs (***Supplementary file 1***). For simplicity in clinical application, we chose the clinical detection outliers of CA-125, CA19-9, CA72-4, AFP, and CEA as the basis for further analysis.

Construction of risk score based on serum CA125, CEA, and AFP levels to predict patient outcomes

To further investigate the role of the platelet indexes in prognosis prediction, the CA125, CEA, and AFP levels were combined to establish the risk score. EC patients included in the study were randomly divided into a test cohort and a validation cohort according to a 1:1 ratio (***Table 3***). By introducing the lasso regression algorithm and the concept of penalty coefficient, the variables of the training cohort

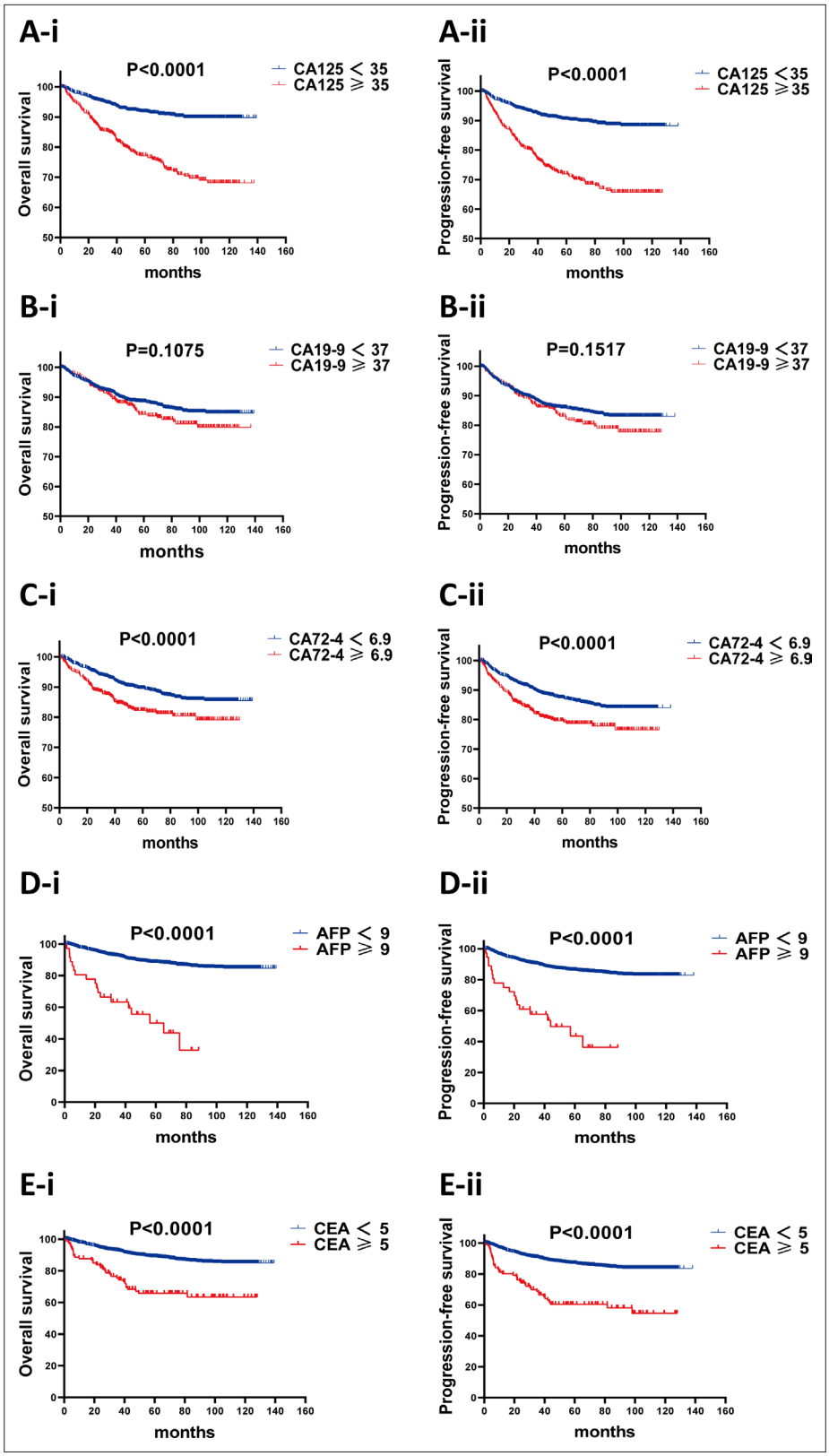

Figure 3. The Kaplan–Meier statistical analysis was used to analyze the overall survival curves and the progression-free survival curves of patients with different levels of each biomarker. (**A–i**) Kaplan–Meier curves of OS for different groups of CA125 level; (**A–ii**) Kaplan–Meier curves of PFS for different groups of CA125 level; (**B–i**) Kaplan–Meier curves of OS for different groups of CA19-9 level; (**B–ii**) Kaplan–Meier curves of PFS for different groups of CA19-9

Figure 3 continued on next page

Figure 3 continued

level; (**C–i**) Kaplan–Meier curves of OS for different groups of CA72-4 level; (**C–ii**) Kaplan–Meier curves of PFS for different groups of CA72-4 level; (**D-i**) Kaplan–Meier curves of OS for different groups of AFP level; (**D-ii**) Kaplan–Meier curves of PFS for different groups of AFP level; (**E–i**) Kaplan–Meier curves of OS for different groups of CEA level; (**E–ii**) Kaplan–Meier curves of PFS for different groups of CEA level.

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Source data 1. Preoperative serum CA125, CA19-9, CA72-4, AFP, and CEA levels, and prognostic data of the endometrial cancer patient population in this study.

Figure supplement 1. The Kaplan–Meier statistical analysis was used to analyze the overall survival curves and the progression-free survival curves of patients with different levels of each biomarker.

are initially screened to solve the collinearity problem that may occur in the process of variable selection and the overfitting problem in the process of model construction. The lasso regression coefficient of OS was plotted (***Figure 4A***). The lasso-cox regression model was used to screen the independent serum tumor markers for OS in EC patients (***Figure 4***). CA125, AFP, and CEA were screened as the best modeling markers to construct the EC prognostic risk score model based on the results of multi-factor Cox regression. The risk score of each patient was calculated according to the regression coefficient and serological level of the three serological tumor markers. Risk score = 0.000382628046162818 × level of CA125 +0.00321692637528874 × level of AFP +0.0269225927239563 × level of CEA. The risk score of each patient was calculated according to the above risk scoring formula, and the ROC curve was drawn (***Figure 4***). The sensitivity and specificity corresponding to different risk probabilities were calculated using the ROC curve in the training cohort, and the Yoden index corresponding to different risk probabilities was further calculated (Yoden index = sensitivity + specificity –1). The maximum Yoden index was used to determine the optimal risk threshold of OS predicted by the model, which was 0.0717. According to this, patients were divided into high and low risk groups, and the risk curves of the two groups were drawn. The death rate of patients in the high risk group was higher, indicating that the high risk group was more likely to have poor prognosis. In the verification cohort, the ROC curve of the drawing time shows that the AUC value of 1–5 years in the verification cohort ranges from 0.740 to 0.808, indicating that the above model still has good prediction performance in the external verification cohort. The risk score of each sample was calculated according to the above risk scoring formula (***Figure 4***), and the Kaplan–Meier survival curve was drawn, indicating that the survival rate of the high-risk rating group was significantly lower than that of the low-risk rating group ($p<0.0001$).

In the entire EC cohort, risk score, age, grade, FIGO stage, pathological classification and other clinical information were integrated into the univariate and multivariate Cox regression analysis of OS and PFS (***Table 4***). Univariate analysis showed that there was a statistical difference in risk scores between OS and PFS ($p<0.0001$). Kaplan–Meier survival curve showed that the survival time of PFS and OS in the high-risk rating group was significantly lower than that in the low-risk rating group ($p<0.0001$; ***Figure 5***). Combined with multivariate Cox regression analysis, RS was an independent prognostic factor for patients with PFS and OS ($p<0.0001$) and could be independently used as a prognostic indicator for patients with EC (***Table 4***).

Nomogram construction and validation

Based on the results of the Cox regression models in ***Table 2***, nomograms were constructed by integrating the independent prognostic factors for PFS and OS (***Figure 6***). In addition to the risk score, FIGO stage, age, and degree of differentiation were included in the nomograms. Summing the scale points corresponding to each variable on the nomogram corresponds to the total score on the bottom scale, which represents the 3- and 5-year PFS and OS probabilities that were predicted by the model. The C-index of the multivariate prognostic model of PFS according to the risk score, age, FIGO stage, and differentiation variables was 0.836 (95% CI=0.822–0.918). The C-index of the multivariate prognostic model of OS according to the risk score, age, FIGO stage, and differentiation variables was 0.837. The C-index showed that the prediction results of the prognostic model had a high probability of agreement with the actual observation results. These findings indicated that the model had good discriminative ability to separate patients with different survival outcomes. The calibration plots of the nomogram show good consistency between the actual observations and the predicted probabilities of PFS and OS (***Figure 7***).

Table 3. Patient biomarkers and characteristics of the training and validation cohorts.

Variables	Training cohort (n=1041)	Validation cohort (n=1040)	Total (n=2081)	p-Values
AFP (U/mL)				0.3145
AFP <9	1020 (98.0)	1025 (98.6)	2045 (98.3)	
AFP ≥9	21 (2.0)	15 (1.4)	36 (1.7)	
CEA (U/mL)				0.4222
CEA <5	992 (95.3)	983 (94.5)	1975 (95.0)	
CEA ≥5	49 (4.7)	57 (4.5)	106 (5.0)	
CA (U/mL)				0.8162
CA125<35	754 (72.5)	758 (72.9)	1512 (72.7)	
CA125≥35	287 (27.5)	282 (27.1)	569 (27.3)	
CA19-9 (U/mL)				0.1844
CA19−9<37	851 (81.8)	873 (83.9)	1724 (82.8)	
CA19−9≥37	190 (18.2)	167 (16.1)	357 (17.2)	
CA72-4 (U/mL)				0.4743
CA72−4<.6.9	781 (75.0)	766 (73.7)	1547 (74.3)	
CA72−4≥.6.9	260 (25.0)	274 (26.3)	534 (25.7)	
Age (years)				0.672
<60	649 (62.3)	639 (61.4)	1288 (61.9)	
≥60	392 (37.7)	401 (38.6)	793 (38.1)	
Menopause				0.2934
no	311 (29.9)	289 (27.8)	600 (28.8)	
yes	730 (70.1)	751 (72.2)	1481 (71.2)	
BMI (kg/m)				0.575
<28	849 (81.6)	858 (82.5)	1707 (82)	
≥28	192 (18.4)	182 (17.5)	374 (18)	
FIGO stage				0.5539
Ⅰ	828 (79.5)	817 (78.5)	1645 (79.1)	
Ⅱ-Ⅳ	213 (20.5)	224 (21.5)	436 (20.9)	
Pathological type				0.5087
Type I	979 (94.0)	985 (94.7)	1964 (94.4)	
Type Ⅱ	62 (6.0)	55 (5.3)	117 (5.6)	
Myometrial invasion				0.7947
<1/2	825 (79.3)	829 (79.7)	1654 (79.5)	
≥1/2	216 (20.7)	211 (20.3)	427 (20.5)	
Differentiation			.	0.7425
G1	530 (50.9)	522 (50.2)	1052 (50.6)	
G2/3	511 (49.1)	518 (49.8)	1029 (49.4)	
LVSI			.	0.2687
Negative	1011 (97.1)	1001 (96.3)	2012 (96.7)	
Positive	30 (2.9)	39 (3.7)	69 (3.3)	

Table 3 continued on next page

Table 3 continued

Variables	Training cohort (n=1041)	Validation cohort (n=1040)	Total (n=2081)	p-Values
Death				0.0884
no	896 (86.1)	921 (88.6)	1817 (87.3)	
yes	145 (13.9)	119 (11.4)	264 (12.7)	

Discussion

EC is one of the most common female malignancies as well as the third most universal female genital malignancy. While most patients are diagnosed relatively early, a significant number of patients are still being diagnosed in the later stages of disease. The FIGO staging system has been guiding significance for the assessment of prognosis and treatment of EC patients; however, prognostic status could vary significantly even for patients divided into the same stage.

In our study, we included 2081 patients with EC undergoing surgical treatment to perform an analysis of independent prognostic factors. Patient groups with high and low levels of CA125, CA199, AFP, CEA, and CA724 were established, high level of CA125 is considered as a risk factor for poor prognosis of EC patients, which is consistent with previous studies. It is worth mentioning that in our patient cohort, AFP and CEA were also identified as independent prognostic factors for EC patients, and patients with high levels of AFP and CEA showed poor prognosis. Therefore, we constructed a new EC risk parameter using AFP, CEA, and CA125 levels to predict the clinical prognosis of PFS and OS in EC patients.

Least absolute shrinkage and selection operator (LASSO) is a regression analysis method, which can simultaneously select and regularize variables to improve the prediction accuracy and interpretability of the generated statistical model. The algorithm has been widely applied to Cox proportional risk regression model (***Reis et al., 2010***; ***Gadducci et al., 2004***), and is used for survival analysis of high-dimensional data. Lasso Cox regression can be used to screen out independent variable sets with strong explanatory power for dependent variables, which is helpful to improve the refinement and accuracy of the prediction model. It is mainly applied to the selection of variables with small samples and multiple factors. This study used LASSO regression model to explore the prognostic factors of EC patients. Based on multiple serum tumor markers, this risk score showed high efficacy in predicting patient outcomes, and patients in the high-risk group assigned by this score showed worse PFS (HR = 3.078 95% CI = 1.806–5.245) and OS (HR = 3.060 95% CI = 1.706–5.490). The risk profile of this score may be a more targeted and powerful prognostic assessment for predicting positive clinical outcomes, and may be a more effective classification tool for EC patients. We further constructed a nomogram using risk score, FIGO staging, grading, and patient age. The c index of the model is greater than 0.8, and the calibration curve is close to the diagonal, indicating that the model has good predictive ability to determine the prognostic value of this combination for EC patients.

Tumor biomarkers are substances expressed and secreted in the process of tumorigenesis and development. They may indicate the existence of new tumor growth, which means relatively large tumor burden and invasive phenotype. Tumor biomarkers can be obtained from a variety of body fluids, such as blood (including serum and plasma), urine, cerebrospinal fluid, sputum, and saliva. Blood is the easiest to sample, and because of the wide contact between blood and organs, blood can collect more valuable biomarkers. Serum tumor biomarkers are widely used in clinic and provide important information for early diagnosis, evaluation of curative effect, prognosis, and prediction of metastasis and recurrence of malignant tumors. The measurement of serum tumor biomarkers is convenient and economical. Therefore, they are promising tools for guiding treatment plans and long-term monitoring of recurrence and metastasis. Among the numerous serum tumor markers, glycoproteins or glycans are widely used in clinical practice because abnormal glycosylation is associated with cancer cell invasion and metastasis (***Reis et al., 2010***). The vast majority of tumor markers currently approved by the U.S. Food and Drug Administration (FDA) in clinical practice are glycoproteins and glycans; the CA125, CA199, AFP, CEA, and CA724 involved in this study are all glycosylated tumor markers that have been applied in clinical practice. The area of research investigating biomarkers for EC is relatively well established (***Gadducci et al., 2004***; ***Staff et al., 2011***; ***Baser et al., 2014***). According to numerous previous studies, it has been demonstrated that the tumor biomarker CA125 is most closely

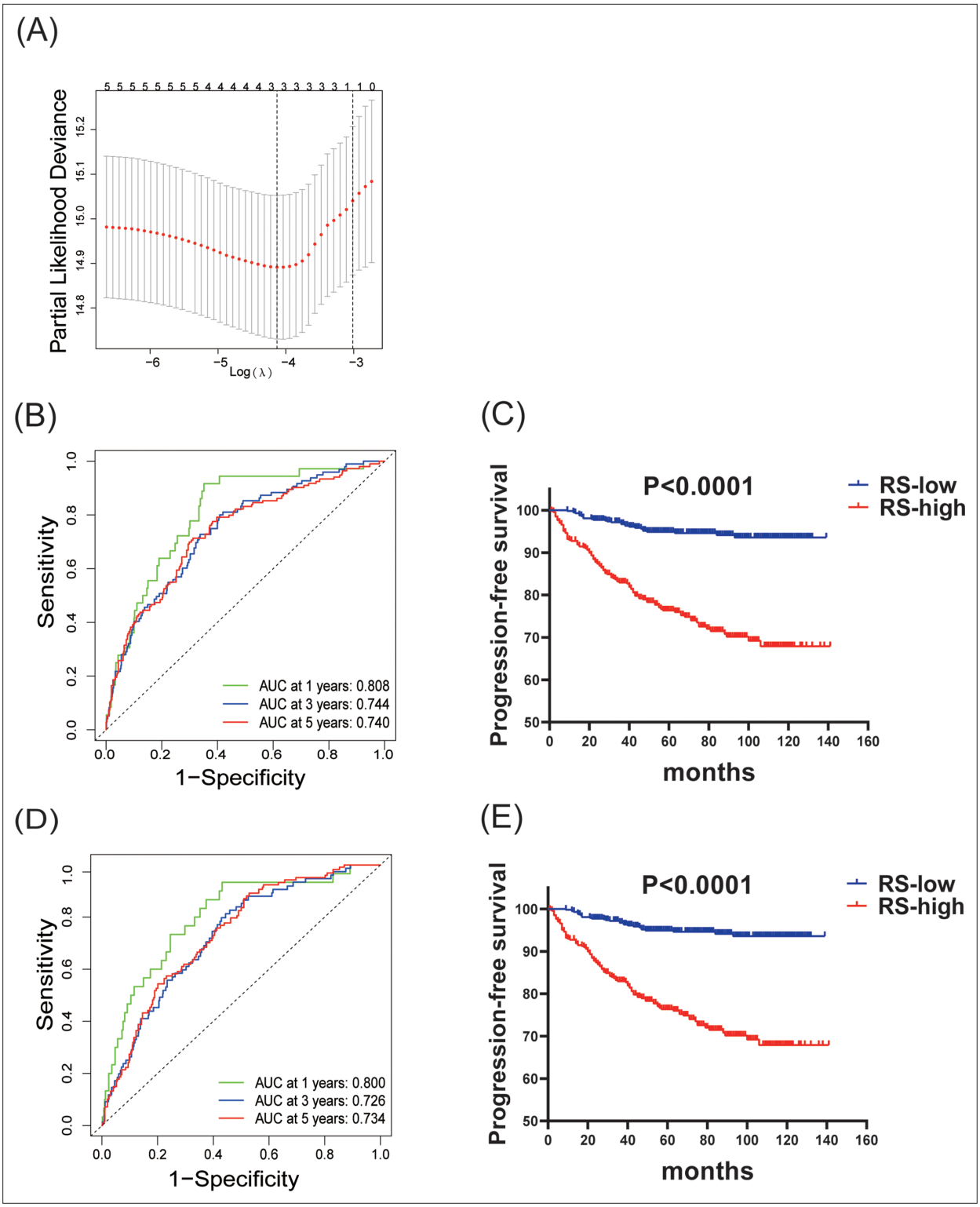

Figure 4. Contribution of Risk-score of OS. (**A**) Screening prognostic factor by Lasso regression analysis, Cross-verify the confidence interval distribution of the variable and the target variable; (**B**) Receiver operating characteristics (ROC) curve analysis of the Risk-score for OS in training cohorts; (**C**) Kaplan–Meier curves of OS for each Risk-score level group in training cohorts; (**D**) Receiver operating characteristics (ROC) curve analysis of the Risk-score for OS in validation cohorts; (**E**) Kaplan–Meier curves of OS for each Risk-score level group in validation cohorts.

The online version of this article includes the following source data for figure 4:

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patient population in this study.

Table 4. Overall survival and progression-free survival of the Risk score with other preoperative biomarkers and clinicopathological variables.

	OS				PFS			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
Variables	p-Value	HR (95% CI)	p-Value	HR (95% CI)	p-Value	HR (95% CI)	p-Value	HR (95% CI)
Risk score	<0.0001		<0.0001		<0.0001		<0.0001	
Low		Ref		Ref		Ref		Ref
High		5.043 (3.754–6.774)		3.060 (1.706–5.490)		5.512 (4.171–7.274)		3.078 (1.806–5.245)
AFP(U/mL)	<0.0001		<0.0001		<0.0001		<0.0001	
AFP <9		Ref		Ref		Ref		Ref
AFP ≥9		6.039 (3.737–9.758)		5.831 (2.872–11.838)		5.559 (3.490–8.855)		5.935 (2.970–11.861)
CEA(U/mL)	<0.0001		0.179		<0.0001		0.357	
CEA <5		Ref		Ref		Ref		Ref
CEA ≥5		3.21 (2.237–4.598)		1.585 (0.81–3.102)		3.549 (2.554–4.932)		1.345 (0.716–2.527)
CA125(U/mL)	<0.0001		0.089		<0.0001		0.007	
CA125<35		Ref		Ref		Ref		Ref
CA125≥35		3.158 (2.480–4.020)		1.480 (0.942–1.326)		3.296 (2.634–4.124)		1.754 (1.164–2.642)
CA19-9(U/mL)	0.108				0.152			
CA199<37		Ref				Ref		
CA199≥37		1.28 (0.948–1.717)				1.226 (0.928–1.619)		
CA72-4(U/mL)	<0.0001		0.779		<0.0001		0.900	
CA724<6.9		Ref		Ref		Ref		Ref
CA724≥6.9		1.701 (1.315–2.201)		1.066 (0.681–1.670)		1.686 (1.329–2.140)		1.026 (0.684–1.541)
Age (years)	<0.0001		0.011		<0.0001		0.054	
<60		Ref		Ref		Ref		Ref
≥60		2.550 (1.996–3.257)		1.833 (1.149–2.922)		2.266 (1.808–2.839)		1.503 (0.993–2.214)
Fertility situation	0.082				0.625			
None								
Yes		1.568 (0.945–2.602)				1.062 (0.834–1.354)		
Menopause	<0.0001		0.282		<0.0001		0.107	
No		Ref		Ref		Ref		Ref
Yes		2.31 (1.649–3.242)		1.442 (0.740–2.813)		2.056 (1.530–2.765)		1.631 (0.900–2.957)
BMI (kg/m^2)	0.140				0.053			
<28		Ref				Ref		
≥28		1.244 (0.931–1.662)				1.301 (0.996–1.698)		
FIGO stage	<0.0001		<0.0001		<0.0001		<0.0001	
Ⅰ		Ref		Ref		Ref		Ref
Ⅱ-Ⅳ		6.214 (4.867–7.934)		2.907 (1.819–4.671)		6.854 (5.458–8.607)		3.526 (2.283–5.446)
Pathological type	<0.0001		0.003		<0.0001		0.003	
Type I		Ref		Ref		Ref		Ref
Type Ⅱ		6.317 (4.628–8.621)		2.330 (1.341–4.047)		6.019 (4.502–8.048)		2.170 (1.304–3.609)

Table 4 continued on next page

Table 4 continued

	OS				PFS			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
Variables	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**	**p-Value**	**HR (95% CI)**
Myometrial invasion	<0.0001		0.605		<0.0001		0.583	
<1/2		Ref		Ref		Ref		Ref
≥1/2		2.982 (2.326–3.824)		1.135 (0.703–1.830)		3.003 (2.385–3.782)		1.128 (0.734–1.733)
Differentiation	<0.0001		0.018		<0.0001		0.002	
G1		Ref		Ref		Ref		Ref
G2/3		3.452 (2.602–4.582)		1.914 (1.120–3.271)		3.453 (2.658–4.485)		2.180 (1.328–3.580)
LVSI	<0.0001		0.005		<0.0001		0.003	
Negative		Ref		Ref		Ref		Ref
Positive		3.674 (2.388–5.652)		2.439 (1.310–4.541)		4.046 (2.752–5.949)		2.358 (1.345–4.135)
Diabetes	0.169				0.068			
No		Ref				Ref		
Yes		1.23 (0.916–1.653)				1.287 (0.981–1.687)		
Hypertension	0.810				0.711			
No		Ref				Ref		
Yes		1.031 (0.802–1.325)				0.957 (0.759–1.207)		
Comorbidities	0.191				0.183			
No		Ref				Ref		
Yes		1.559 (0.802–3.030)				1.535 (0.817–2.882)		
Expression of ER	<0.0001		0.874		<0.0001		0.870	
Negative		Ref		Ref		Ref		Ref
Positive		0.682 (0.556–0.836)		1.031 (0.710–1.496)		0.637 (0.530–0.767)		1.029 (0.727–1.456)
Expression of PR	<0.0001		0.265		<0.0001		0.440	
Negative		Ref		Ref		Ref		Ref
Positive		0.569 (0.469–0.690)		0.804 (0.548–1.180)		0.559 (0.468–0.667)		0.871 (0.613–1.237)
Expression of P53	0.006		0.799		<0.0001		0.482	
Negative		Ref		Ref		Ref		Ref
Positive		1.515 (1.129–2.033)		0.962 (0.713–1.297)		1.732 (1.323–2.267)		1.103 (0.839–1.450)

related to ***Niloff et al., 1984*** first reported that serum CA125 was elevated in women with recurrent and advanced EC in 1984. Subsequent studies have consistently shown that there is a correlation between serum CA125 concentration and the clinicopathological parameters and outcomes of poor EC (***Kurihara et al., 1998***; ***Sood et al., 1997***; ***Han et al., 2010***; ***Coll-de la Rubia et al., 2020***). CA125 can also be used as a preoperative indicator to evaluate whether regional lymph nodes are invaded, which can predict OS and DFS. In these studies, CA125 is mainly used to monitor the postoperative recovery and long-term prognosis of EC patients (***Sood et al., 1997***; ***Kotowicz et al., 2017***; ***Nithin et al., 2018***), or to improve the EC positive rate in combination with HE4. In this study, both univariate and multivariate analyses showed that elevated serum levels of CA125 indicated a poor prognosis for OS and PFS in EC patients. Patients with serum concentration of CA125 ≥35.0 mm had a 1.9-fold higher risk of death due to cancer compared to patients with baseline levels of CA125 (p=0.003, HR = 1.989, 95% CI: 1.270–3.116). This is consistent with previously published literature.

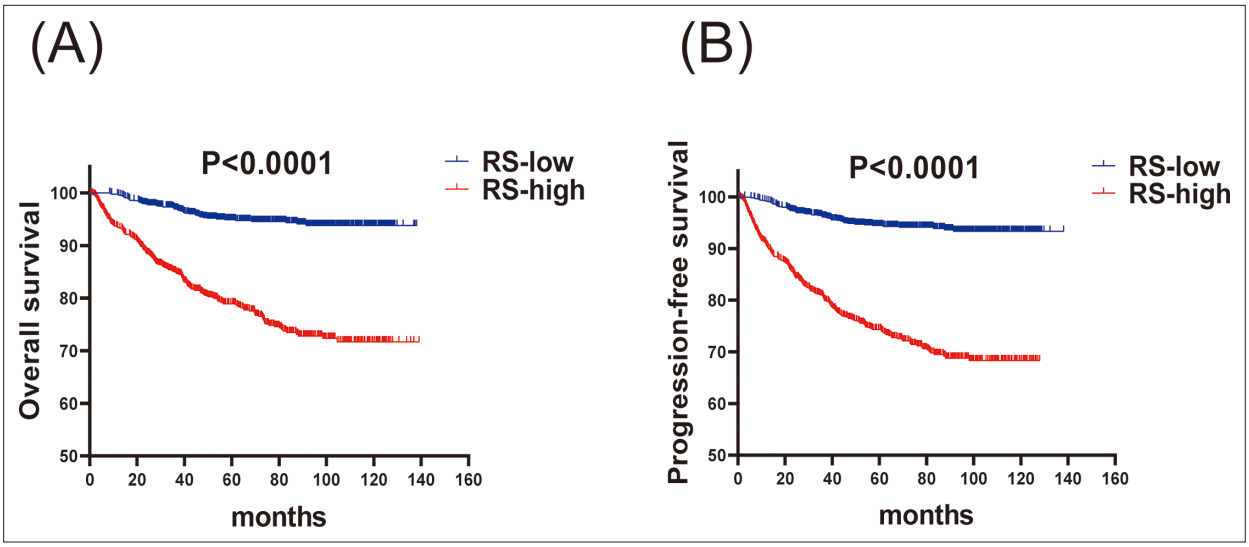

Figure 5. Verification of Risk-score in all endometrial cancer (EC) patients. (**A**) Kaplan–Meier curves of OS for each Risk-score level group in all EC patients; (**B**) Kaplan–Meier curves of PFS for each Risk-score level group in all EC patients.

The online version of this article includes the following source data for figure 5:

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patient population in this study.

CEA is an acidic glycoprotein with the characteristics of human embryonic antigen. It exists on the surface of cancer cells differentiated from endodermal cells and belongs to the structural protein of cell membrane. It is synthesized in the small intestine, liver, and pancreas during the embryonic stage, is a fetal glycoprotein with extremely low levels of CEA in human serum after birth, and elevated CEA in adults usually predicts the tumorigenesis. CEA was first used as a specific marker for the diagnosis of colorectal cancer (***Lin et al., 2018***), and CEA can also be used as a good indicator for the prognosis evaluation of colorectal cancer. In addition to elevated levels in colorectal cancer, ***Hao et al., 2019*** found that in 42 different diseases, serum CEA levels of patients were significantly higher than those in the healthy population. In recent years, research shows that CEA has certain reference value in the diagnosis of female reproductive system tumors (***Zhou et al., 2021***). A study by ***Gadducci et al., 2018*** shows that CEA is also overexpressed in elderly EC patients. Although it cannot be used as a specific indicator of elderly EC, it plays an important role in the diagnosis of differentiation of endometrial tumor and evaluation of curative effects. A recent study reported that pre-treatment CEA and CA125 levels could be used to predict lymph node metastasis in EC, and the investigators developed a cost-effective method to distinguish EC patients with higher and lower risk of lymph node metastasis (***Huang et al., 2023***). A multicenter retrospective cohort study included 1559 EC patients in FIGO stage I; CEA was used to construct the nomograms of prognostic prediction with good accuracy and consistency (***Lin et al., 2024***). Unlike the previous study, this study included all EC patients in FIGO stage I–IV, focusing on the potential use of serum CEA as a prognostic tool for EC survival, confirmed that patients with elevated serum CEA concentrations had a worse prognosis than those with normal serum CEA concentrations, and established new risk scores using CEA, as well as valuable survival prediction models.

AFP is a glycoprotein that is mainly synthesized by fetal liver cells and the yolk sac. It is found in high concentrations in the fetal circulatory system and declines shortly after birth to be functionally replaced by albumin. It is present in extremely low concentrations in adult serum (***Gitlin et al., 1972***). AFP has many important physiological functions, including transport functions, bidirectional regulation of growth, immunosuppression, and is also involved in T cell-induced apoptosis. AFP is the most widely used serological marker for the diagnosis of hepatocellular carcinoma (***Liu et al., 2022***). It is also elevated in a variety of extrahepatic tumors, including yolk sac tumors, gastrointestinal tumors, and pancreatic, gallbladder, lung, and bladder cancers (***Akazawa et al., 2018***; ***Suzuki et al., 2021***; ***Yamazawa et al., 2017***). Studies have found that AFP-producing endometrial carcinoma is a clinicopathologic subtype of endometrial carcinoma (***Li et al., 2018***; ***Yin et al., 2019***; ***Liu et al., 2021***), which

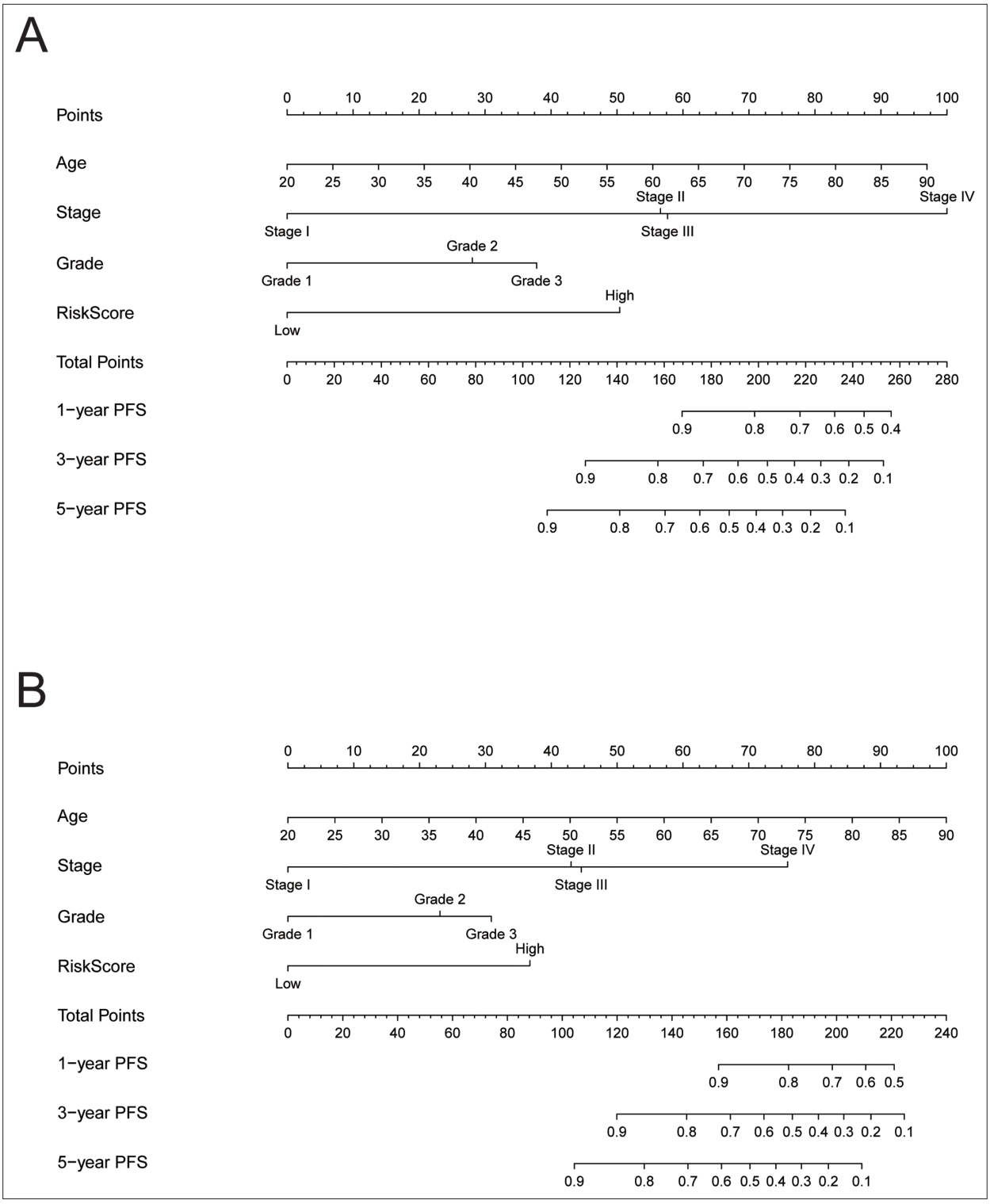

Figure 6. Nomograms for predicting the postoperative PFS (**A**) and OS (**B**) for 1, 3 and 5 years in endometrial patients.

The online version of this article includes the following source data for figure 6:

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patient population in this study.

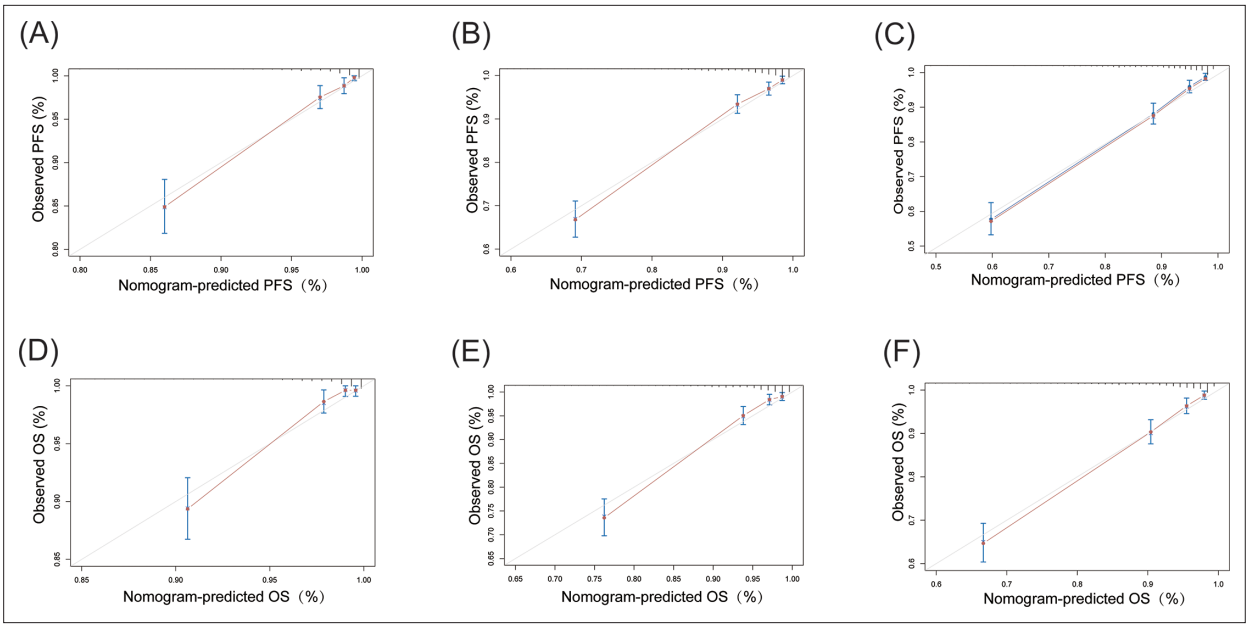

Figure 7. Calibration plot of the nomograms of PFS and OS in endometrial cancer (EC) patients. (**A**) Calibration plots of the nomograms of PFS for 1 year in EC patients; (**B**) Calibration plots of the nomograms of PFS for 3 years in EC patients; (**C**) Calibration plots of the nomograms of PFS for 5 years in EC patients; (**D**) Calibration plots of the nomograms of OS for 1 year in EC patients; (**E**) Calibration plots of the nomograms of OS for 3 years in EC patients; (**F**) Calibration plots of the nomograms of OS for 5 years in EC patients.

The online version of this article includes the following source data for figure 7:

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patient population in this study.

occurs in 1.8% of endometrial carcinoma cases (***Otani et al., 2021***). This classification is related to TP53 abnormality and vascular invasion. Even in FIGO stage I cases, these cancers can show aggressive behavior (***Liu et al., 2021***; ***Otani et al., 2021***). Similarly, in our study cohort, 1.73% of patients showed AFP expression above the normal range, and this part of patients also showed significant adverse prognosis. We identified serum concentrations of AFP ≥9 ng/ml in EC patients as a marker for poor prognosis for PFS ($p<0.0001$, HR = 7.437, 95% CI: 3.717–14.879) as well as OS ($p<0.0001$, HR = 6.914, 95% CI: 3.374–14.165). There are few studies on the association between AFP and EC, mostly case-reported rather than cohort studies. This study provides the description of serum AFP can be used as an independent risk factor for the prognosis of EC through a large-sample cohort study. AFP measurement may provide further information after surgical risk stratification to help clinicians predict prognosis and provide a basis for clinical treatment.

The study still has limitations. This study is a single-center cohort study and the conclusions may not be generalized to the wider population. Some patient accurate information was unable to be obtained due to their loss of follow-up. However, as a prestigious and important medical institution, the research center has gathered a large number of patients in the region, a large EC patient sample population was enrolled, and the data from the electronic medical record system of Shengjing Hospital of China Medical University were verified, this provides a firm basis for the conclusions drawn. We validated the prognostic value of this risk score in patients with FIGO stage I (***Supplementary files 2 and 4***) and stages II-IV (***Supplementary files 4 and 5***) through stratified analysis. Our findings revealed that the risk score predicted OS and PFS with greater accuracy in the FIGO stage I population compared to stages II-IV. This discrepancy may be attributed to the fact that over 70% of our patients were in stage I, both in our study and in clinical practice. The characteristics of patients with stage II–IV EC differ from those of the overall patient population, and the higher mortality rate in these stages contributed to the bias in the model. Future studies should include multicenter patient data to further validate our results and improve data quality, ensuring the accuracy of our findings. In conclusion, CA125, AFP, and CEA concentrations are promising biomarkers for EC and a valuable independent factor for predicting patient prognosis and survival. Furthermore, a risk score created by

combining CA125, AFP, and CEA levels to predict the prognosis of EC was found to be highly accurate and valuable for clinical use.

Additional information

Funding

Funder	Grant reference number	Author
National Key R&D Program of China	2022YFC2704400	Xiao-xin Ma
National Natural Science Foundation of China	81872123	Xiao-xin Ma

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Zi-hao Wang, Software, Validation, Writing – original draft, Writing – review and editing; Qing Li, Jiahui Gu, Conceptualization, Formal analysis; Yun-zheng Zhang, Conceptualization, Software; Shu-wen Ge, Investigation, Methodology; Lu-he Shan, Investigation, Methodology, Writing – review and editing; Bo Wang, Software, Methodology; Zi-yu Zhang, Data curation, Investigation, Methodology; Qi-jun Wu, Formal analysis, Validation, Methodology, Writing – review and editing; Hui Yang, Supervision, Validation; Xiao-xin Ma, Conceptualization, Resources, Supervision, Funding acquisition, Writing – review and editing

Author ORCIDs

Jiahui Gu https://orcid.org/0000-0001-8080-7258
Bo Wang https://orcid.org/0000-0001-5423-9111
Qi-jun Wu https://orcid.org/0000-0001-9421-5114
Xiao-xin Ma https://orcid.org/0000-0003-0271-8035

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.10.7554/eLife.94480.sa1

Additional files

Supplementary files

Supplementary file 1. Overall survival and progression-free survival of the preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables.

Supplementary file 2. Overall survival and progression-free survival of the preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables in Stage I EC patients.

Supplementary file 3. Overall survival and progression-free survival of the preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables in Stage II–IV EC patients.

Supplementary file 4. Overall survival and progression-free survival of the Risk score and preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables in Stage I EC patients.

Supplementary file 5. Overall survival and progression-free survival of the Risk score and preoperative CA125, CA19-9, CA72-4, AFP, and CEA with other clinicopathological variables in Stage II–IV EC patients.

MDAR checklist

Source data 1. All clinical baseline data, serological indicators, and prognostic data of the endometrial cancer patience population.

Data availability

There are no linked research data sets for this paper. We have uploaded all data as the resource file.

References

Abbink K, Zusterzeel PL, Geurts-Moespot AJ, van Herwaarden AE, Pijnenborg JM, Sweep FC, Massuger LF. 2018. HE4 is superior to CA125 in the detection of recurrent disease in high-risk endometrial cancer patients. *Tumour Biology* **40**:1010428318757103. DOI: https://doi.org/10.1177/1010428318757103, PMID: 29463191

Akazawa Y, Saito T, Hayashi T, Yanai Y, Tsuyama S, Akaike K, Suehara Y, Takahashi F, Takamochi K, Ueyama H, Murakami T, Watanabe S, Nagahara A, Yao T. 2018. Next-generation sequencing analysis for gastric adenocarcinoma with enteroblastic differentiation: emphasis on the relationship with hepatoid adenocarcinoma. *Human Pathology* **78**:79–88. DOI: https://doi.org/10.1016/j.humpath.2018.04.022, PMID: 29751042

Angioli R, Plotti F, Capriglione S, Montera R, Damiani P, Ricciardi R, Aloisi A, Luvero D, Cafà EV, Dugo N, Angelucci M, Benedetti-Panici P. 2013. The role of novel biomarker HE4 in endometrial cancer: a case control prospective study. *Tumor Biology* **34**:571–576. DOI: https://doi.org/10.1007/s13277-012-0583-0, PMID: 23179397

Baser E, Gungor T, Togrul C, Turkoglu O, Celen S. 2014. Preoperative prediction of poor prognostic parameters and adjuvant treatment in women with pure endometrioid type endometrial cancer: what is the significance of tumor markers? *European Journal of Gynaecological Oncology* **35**:513–518. DOI: https://doi.org/10.12892/ejgo24702014, PMID: 25423695

Behrouzi R, Barr CE, Crosbie EJ. 2021. HE4 as a biomarker for endometrial Cancer. *Cancers* **13**:4764. DOI: https://doi.org/10.3390/cancers13194764, PMID: 34638250

Bian J, Sun X, Li B, Ming L. 2017. Clinical significance of serum HE4, CA125, CA724, and CA19-9 in patients with endometrial cancer. *Technology in Cancer Research & Treatment* **16**:435–439. DOI: https://doi.org/10.1177/1533034616666644, PMID: 27562869

Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. 2016. Cancer statistics in China, 2015. *CA* **66**:115–132. DOI: https://doi.org/10.3322/caac.21338, PMID: 26808342

Cheng Y, Yang L, Shi G, Chen P, Li L, Fang H, Chen C. 2021. Ninjurin 2 rs118050317 gene polymorphism and endometrial cancer risk. *Cancer Cell International* **21**:1. DOI: https://doi.org/10.1186/s12935-020-01646-5, PMID: 33397383

Coll-de la Rubia E, Martinez-Garcia E, Dittmar G, Gil-Moreno A, Cabrera S, Colas E. 2020. Prognostic biomarkers in endometrial cancer: a systematic review and meta-analysis. *Journal of Clinical Medicine* **9**:1900. DOI: https://doi.org/10.3390/jcm9061900, PMID: 32560580

Crosbie EJ, Kitson SJ, McAlpine JN, Mukhopadhyay A, Powell ME, Singh N. 2022. Endometrial cancer. *Lancet* **399**:1412–1428. DOI: https://doi.org/10.1016/S0140-6736(22)00323-3, PMID: 35397864

Dochez V, Caillon H, Vaucel E, Dimet J, Winer N, Ducarme G. 2019. Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. *Journal of Ovarian Research* **12**:28. DOI: https://doi.org/10.1186/s13048-019-0503-7, PMID: 30917847

Dotters DJ. 2000. Preoperative CA 125 in endometrial cancer: Is it useful? *American Journal of Obstetrics and Gynecology* **182**:1328–1334. DOI: https://doi.org/10.1067/mob.2000.106251, PMID: 10871446

Espiau Romera A, Coronado Martín PJ, Chóliz Ezquerro M, Cuesta Guardiola T, Adiego Calvo I, Baquedano Mainar L. 2021. Value of preoperative HE4 as predictor of advanced disease in endometrioid endometrial cancer. *International Journal of Gynecology & Obstetrics* **153**:64–70. DOI: https://doi.org/10.1002/ijgo.13473, PMID: 33156532

Felder M, Kapur A, Gonzalez-Bosquet J, Horibata S, Heintz J, Albrecht R, Fass L, Kaur J, Hu K, Shojaei H, Whelan RJ, Patankar MS. 2014. MUC16 (CA125): tumor biomarker to cancer therapy, a work in progress. *Molecular Cancer* **13**:129. DOI: https://doi.org/10.1186/1476-4598-13-129, PMID: 24886523

Gadducci A, Cosio S, Carpi A, Nicolini A, Genazzani AR. 2004. Serum tumor markers in the management of ovarian, endometrial and cervical cancer. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie* **58**:24–38. DOI: https://doi.org/10.1016/j.biopha.2003.11.003, PMID: 14739059

Gadducci A, Guerrieri ME, Cosio S, Fabrini MG, Laliscia C, Attianese D, Rossi A, Ferrero A. 2018. Rates, sites and times of recurrence and clinical outcome of endometrial cancer patients with histologically-positive nodes: An Italian two-center retrospective study. *Anticancer Research* **38**:1695–1703. DOI: https://doi.org/10.21873/anticanres.12403, PMID: 29491104

Gitlin D, Perricelli A, Gitlin GM. 1972. Synthesis of -fetoprotein by liver, yolk sac, and gastrointestinal tract of the human conceptus. *Cancer Research* **32**:979–982 PMID: 4111729.

Han S-S, Lee SH, Kim DH, Kim JW, Park N-H, Kang S-B, Song YS. 2010. Evaluation of preoperative criteria used to predict lymph node metastasis in endometrial cancer. *Acta Obstetricia et Gynecologica Scandinavica* **89**:168–174. DOI: https://doi.org/10.3109/00016340903370114, PMID: 19916890

Hao C, Zhang G, Zhang L. 2019. Serum CEA levels in 49 different types of cancer and noncancer diseases. *Progress in Molecular Biology and Translational Science* **162**:213–227. DOI: https://doi.org/10.1016/bs.pmbts.2018.12.011, PMID: 30905451

Huang S-Y, Fu H-C, Wu C-H, Wang S-C, Ou Y-C, Tsai C-C, Chen Y-Y, Huang S-W, Wang Y-W, Lin H. 2023. Pretreatment carcinoembryonic antigen combined with cancer antigen-125 for predicting lymph node metastasis in endometrial carcinoma: a retrospective cohort study. *Journal of Cancer Research and Clinical Oncology* **149**:11807–11813. DOI: https://doi.org/10.1007/s00432-023-05087-1, PMID: 37405474

Jiang T, Huang L, Zhang S. 2015. Preoperative serum CA125: a useful marker for surgical management of endometrial cancer. *BMC Cancer* **15**:396. DOI: https://doi.org/10.1186/s12885-015-1260-7, PMID: 25964114

Kotowicz B, Fuksiewicz M, Jonska-Gmyrek J, Wagrodzki M, Kowalska M. 2017. Preoperative serum levels of YKL 40 and CA125 as a prognostic indicators in patients with endometrial cancer. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* **215**:141–147. DOI: https://doi.org/10.1016/j.ejogrb.2017.06.021, PMID: 28624692

Kozakiewicz B, Chądzyńska M, Dmoch-Gajzlerska E, Stefaniak M. 2016. Monitoring the treatment outcome in endometrial cancer patients by CEA and TATI. *Tumour Biology* **37**:9367–9374. DOI: https://doi.org/10.1007/s13277-016-4784-9, PMID: 26779635

Kurihara T, Mizunuma H, Obara M, Andoh K, Ibuki Y, Nishimura T. 1998. Determination of a normal level of serum CA125 in postmenopausal women as a tool for preoperative evaluation and postoperative surveillance of endometrial carcinoma. *Gynecologic Oncology* **69**:192–196. DOI: https://doi.org/10.1006/gyno.1998.5018, PMID: 9648586

Laengsri V, Kerdpin U, Plabplueng C, Treeratanapiboon L, Nuchnoi P. 2018. Cervical cancer markers: Epigenetics and microRNAs. *Laboratory Medicine* **49**:97–111. DOI: https://doi.org/10.1093/labmed/lmx080, PMID: 29378033

Li JJX, Lee JHS, Chan VTC, Yu MY. 2018. Uterine Carcinosarcoma with alpha-fetoprotein-producing hepatoid component: A case report and literature review. *Case Reports in Pathology* **2018**:3972353. DOI: https://doi.org/10.1155/2018/3972353, PMID: 29992073

Lin W, Cao D, Shen K. 2018. Prognostic significance of preoperative serum CEA in primary mucinous ovarian carcinoma: A retrospective cohort study. *Cancer Management and Research* **10**:6913–6920. DOI: https://doi.org/10.2147/CMAR.S186258, PMID: 30588097

Lin H, Wu CH, Ou YC, Huang SW, Fu HC. 2024. Integration of pretreatment tumor markers in a nomogram model for prognostic prediction of FIGO stage I endometrial cancer: A multi-institutional cohort study. *International Journal of Gynaecology and Obstetrics* **165**:1244–1256. DOI: https://doi.org/10.1002/ijgo.15362, PMID: 38287783

Liu Y, Zhou R, Wang S, Zhang G. 2021. Extra-hepatic hepatoid carcinomas in female reproductive system: Three case-reports with a literature review. *Cancer Management and Research* **13**:1625–1636. DOI: https://doi.org/10.2147/CMAR.S288913, PMID: 33628053

Liu S, Sun L, Yao L, Zhu H, Diao Y, Wang M, Xing H, Lau WY, Guan M, Pawlik TM, Shen F, Xu M, Tong X, Yang T. 2022. Diagnostic performance of AFP, AFP-L3, or PIVKA-II for hepatitis C virus-associated hepatocellular carcinoma: A multicenter analysis. *Journal of Clinical Medicine* **11**:5075. DOI: https://doi.org/10.3390/jcm11175075, PMID: 36079006

Makker V, MacKay H, Ray-Coquard I, Levine DA, Westin SN, Aoki D, Oaknin A. 2021. Endometrial cancer. *Nature Reviews. Disease Primers* **7**:88. DOI: https://doi.org/10.1038/s41572-021-00324-8, PMID: 34887451

Miller DS, King LP. 2008. Gynecologic oncology group trials in uterine corpus malignancies: recent progress. *Journal of Gynecologic Oncology* **19**:218–222. DOI: https://doi.org/10.3802/jgo.2008.19.4.218, PMID: 19471650

Moore RG, Miller CM, Brown AK, Robison K, Steinhoff M, Lambert-Messerlian G. 2011. Utility of tumor marker HE4 to predict depth of myometrial invasion in endometrioid adenocarcinoma of the uterus. *International Journal of Gynecological Cancer* **21**:1185–1190. DOI: https://doi.org/10.1097/IGC.0b013e3182229ad8, PMID: 21720250

Nicklin J, Janda M, Gebski V, Jobling T, Land R, Manolitsas T, McCartney A, Nascimento M, Perrin L, Baker JF, Obermair A, LACE Trial Investigators. 2012. The utility of serum CA-125 in predicting extra-uterine disease in apparent early-stage endometrial cancer. *International Journal of Cancer* **131**:885–890. DOI: https://doi.org/10.1002/ijc.26433, PMID: 21918977

Niloff JM, Klug TL, Schaetzl E, Zurawski VR Jr, Knapp RC, Bast RC Jr. 1984. Elevation of serum CA125 in carcinomas of the fallopian tube, endometrium, and endocervix. *American Journal of Obstetrics and Gynecology* **148**:1057–1058. DOI: https://doi.org/10.1016/s0002-9378(84)90444-7, PMID: 6201072

Nithin KU, Sridhar MG, Srilatha K, Habebullah S. 2018. CA 125 is a better marker to differentiate endometrial cancer and abnormal uterine bleeding. *African Health Sciences* **18**:972–978. DOI: https://doi.org/10.4314/ahs.v18i4.17, PMID: 30766562

Otani T, Murakami K, Shiraishi N, Hagiyama M, Satou T, Matsuki M, Matsumura N, Ito A. 2021. α-Fetoprotein-producing endometrial carcinoma is associated with fetal gut-like and/or hepatoid morphology, lymphovascular infiltration, *TP53* abnormalities, and poor prognosis: Five cases and literature review. *Frontiers in Medicine* **8**:799163. DOI: https://doi.org/10.3389/fmed.2021.799163, PMID: 34977100

Reis CA, Osorio H, Silva L, Gomes C, David L. 2010. Alterations in glycosylation as biomarkers for cancer detection. *Journal of Clinical Pathology* **63**:322–329. DOI: https://doi.org/10.1136/jcp.2009.071035, PMID: 20354203

Siegel RL, Miller KD, Fuchs HE, Jemal A. 2021. Cancer Statistics, 2021. *CA* **71**:7–33. DOI: https://doi.org/10.3322/caac.21654, PMID: 33433946

Sood AK, Buller RE, Burger RA, Dawson JD, Sorosky JI, Berman M. 1997. Value of preoperative CA 125 level in the management of uterine cancer and prediction of clinical outcome. *Obstetrics and Gynecology* **90**:441–447. DOI: https://doi.org/10.1016/s0029-7844(97)00286-x, PMID: 9277659

Staff AC, Trovik J, Eriksson AGZ, Wik E, Wollert KC, Kempf T, Salvesen HB. 2011. Elevated plasma growth differentiation factor-15 correlates with lymph node metastases and poor survival in endometrial cancer. *Clinical Cancer Research* **17**:4825–4833. DOI: https://doi.org/10.1158/1078-0432.CCR-11-0715, PMID: 21616994

Suzuki M, Kasajima R, Yokose T, Ito H, Shimizu E, Hatakeyama S, Yokoyama K, Yamaguchi R, Furukawa Y, Miyano S, Imoto S, Yoshioka E, Washimi K, Okubo Y, Kawachi K, Sato S, Miyagi Y. 2021. Comprehensive molecular analysis of genomic profiles and PD-L1 expression in lung adenocarcinoma with a high-grade fetal adenocarcinoma component. *Translational Lung Cancer Research* **10**:1292–1304. DOI: https://doi.org/10.21037/tlcr-20-1158, PMID: 33889510

Takahashi A, Matsuura M, Matoda M, Nomura H, Okamoto S, Kanao H, Kondo E, Omatsu K, Kato K, Utsugi K, Takeshima N. 2017. Clinicopathological features of early and late recurrence of endometrial carcinoma after surgical resection. *International Journal of Gynecological Cancer* **27**:967–972. DOI: https://doi.org/10.1097/IGC.0000000000000984, PMID: 28486241

Virji MA, Mercer DW, Herberman RB. 1988. Tumor markers in cancer diagnosis and prognosis. *CA* **38**:104–126. DOI: https://doi.org/10.3322/canjclin.38.2.104, PMID: 3127000

Yamazawa S, Ushiku T, Shinozaki-Ushiku A, Hayashi A, Iwasaki A, Abe H, Tagashira A, Yamashita H, Seto Y, Aburatani H, Fukayama M. 2017. Gastric cancer with primitive enterocyte phenotype: An aggressive subgroup of intestinal-type adenocarcinoma. *The American Journal of Surgical Pathology* **41**:989–997. DOI: https://doi.org/10.1097/PAS.0000000000000869, PMID: 28505005

Yin H, Hu Y, Xiu Y, Luo R, Shi H. 2019. A rare case of primary endometrial Hepatoid Adenocarcinoma on 18F-FDG PET/CT. *Clinical Nuclear Medicine* **44**:e425–e427. DOI: https://doi.org/10.1097/RLU.0000000000002538, PMID: 30985410

Zhou L, Wang W, Wang F, Yang S, Hu J, Lu B, Pan Z, Ma Y, Zheng M, Zhou L, Lei S, Song P, Liu P, Lu W, Lu Y. 2021. Plasma-derived exosomal miR-15a-5p as a promising diagnostic biomarker for early detection of endometrial carcinoma. *Molecular Cancer* **20**:57. DOI: https://doi.org/10.1186/s12943-021-01352-4, PMID: 33781255