Am J Cancer Res 2023;13(10):4508-4530
www.ajcr.us /ISSN:2156-6976/ajcr0152300

Original Article

A novel web-based prognostic nomogram and the features influencing the curative effect of chemotherapy and radiotherapy for Paget's disease with invasive ductal carcinoma

Jingkun Qu[1*], Chaofan Li[1*], Mengjie Liu[1], Yusheng Wang[2], Shiyu Sun[1], Zeyao Feng[1], Fei Wu[1], Shuqun Zhang[1], Xixi Zhao[3]

[1]Department of Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 West Fifth Street, Xi'an, Shaanxi, P. R. China; [2]Department of Otolaryngology, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 West Fifth Street, Xi'an, Shaanxi, P. R. China; [3]Department of Radiation Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 West Fifth Street, Xi'an, Shaanxi, P. R. China. [*]Equal contributors.

Received July 18, 2023; Accepted October 9, 2023; Epub October 15, 2023; Published October 30, 2023

Abstract: Paget's disease (PD) of the breast is a rare underlying malignant tumor. Approximately 50% to 60% of patients with mammary PD are concurrently diagnosed with invasive ductal carcinoma (PD-IDC), a condition associated with a worse prognosis than IDC without PD. Thus far, there has been a lack of an accurate and efficient prognostic model for PD-IDC, and the factors influencing the effectiveness of chemotherapy and radiotherapy for these patients remain unknown. In this study, we developed a web-based nomogram based on the data from the Surveillance Epidemiology and End Results (SEER) database. We subjected the model to a series of validation methods, including area under the curve (AUC) values, receiver operating characteristic curve (ROC) analysis, calibration curves, and decision curve analysis (DCA). Our results demonstrated that our model exhibited high discrimination, accuracy, and clinical applicability in predicting the overall survival (OS) of patients with PD-IDC (testing set: three- and five-year AUCs, 0.831 and 0.841, respectively). To further validate our nomogram, we used external data from both our institution and sister hospitals (external data: three- and five-year AUCs, 0.892 and 0.914, respectively). Multivariable Cox regression analysis identified several independent unfavorable prognostic factors for the OS of patients with PD-IDC, including increasing age, high grade, widowed status, higher T stages, and the presence of bone metastases. Furthermore, propensity score matching (PSM)-adjusted analysis was conducted, revealing that chemotherapy did not significantly improve the survival of patients with PD-IDC across molecular subtypes, except for those in the grade III/IV group, where it improved both OS and breast cancer-specific survival (BCSS). Additionally, our findings indicated that only patients with PD-IDC with T4 and N3 stages benefited from radiotherapy, leading to improvements in both OS and BCSS. In conclusion, we have comprehensively analyzed the clinical characteristics and prognosis of patients with PD-IDC, culminating in the development of a user-friendly web-based nomogram for predicting their survival. Our predictive model is not only highly accurate but also offers simplicity, making it accessible for healthcare providers and patients. Furthermore, our stratified analysis highlights that the pathological grade, rather than the molecular subtype, plays a pivotal role in determining the efficacy of chemotherapy in improving the prognosis for patients with PD-IDC, while radiotherapy confers survival benefits to patients with PD-IDC in T4 and N3 stages.

Keywords: Breast cancer, Paget's disease, invasive ductal carcinoma, SEER, nomogram, therapy

Introduction

Paget's disease (PD) of the breast is a rare form of mammary carcinoma, with an estimated incidence of 1% to 3% among all patients with breast cancer [1-3]. Clinically, it is distinguished by the presence of an eczematoid lesion on the nipple, which can extend to the areola, accompanied by other skin signs such as pruritus, erythema, nipple erosion or ulceration, bloody nipple discharge, and nipple retraction [3, 4]. The main histopathological characteristic of mam-

mary PD is the infiltration of large epidermal adenocarcinoma cells, called Paget's cells, within the nipple epidermis [5]. Since mammary PD was first described and connected with an underlying cancer by Sir James Paget in 1874 [6], there has been ongoing debate regarding the prognosis and management of this disease [3].

PD of the breast can be categorized into three groups based on the underlying malignancy, and this categorization is associated with the prognosis of the disease. Approximately 50% to 60% of patients with mammary PD are found to have concurrent invasive ductal carcinoma (PD-IDC), while 30% to 40% of patients also exhibit ductal carcinoma in situ (PD-DCIS). Only 10% of patients present with skin changes alone (PD) [7]. Both PD with IDC or DCIS are associated with worse survival outcomes and more aggressive tumor characteristics compared to the respective diseases without PD [4, 8-10]. Nevertheless, for patients with PD-DCIS treated with breast-conserving surgery (BCS) and radiotherapy, the 15-year overall survival (OS) and breast cancer-specific survival (BCSS) rates can reach as high as 90% and 97%, respectively [11]. However, the prognosis of PD-IDC is much worse, with > 50% of patients with PD-IDC testing positive for lymph node metastasis, compared to approximately 30% of patients with IDC alone [7]. The five-year OS of patients with PD-IDC with positive lymph nodes is merely 20% to 25% [12]. Considering this particularity, accurate prognostic prediction becomes pivotal for these patients. Unfortunately, previous prognostic models [13-15] designed for IDC failed to consider the differences between PD-IDC and IDC, as they excluded PD-IDC from their considerations. Consequently, these models resulted in inaccurate assessments for patients with PD-IDC. Moreover, these models used traditional nomograms, which were not user-friendly for healthcare providers and patients and could only predict outcomes at specific time points. Thus, the need arises to construct a novel web-based prognostic model specifically tailored for patients with PD-IDC.

Moreover, PD is a rare disease, making it challenging to conduct randomized controlled trials to compare the effects of different treatments. Consequently, we are often limited to investigating it through retrospective studies. However, while the majority of retrospective studies have predominantly focused on evaluating the impact of surgery on prognosis [16-21], no one has analyzed the factors influencing the effectiveness of chemotherapy and radiotherapy, highlighting the pressing need for further inquiry.

In our study, we investigated the clinical features and prognostic factors of patients with PD-IDC using the most recent Surveillance Epidemiology and End Results (SEER) database. Our study stands as the first to develop a highly accurate web-based nomogram for predicting the OS of patients with PD-IDC. We also ventured into uncharted territory by exploring the roles of time elapsed from diagnosis and family income in patients with PD-IDC. Additionally, following propensity score matching (PSM), we conducted a novel stratified analysis of chemotherapy and radiotherapy, which had not been reported previously. Our findings revealed that chemotherapy can enhance the survival of patients with PD-IDC with grades III/IV, with no discernible impact from different molecular subtypes. Furthermore, we discovered that only patients with T4 and N3 stages experienced a significant survival improvement following radiotherapy. These findings underscore the significant importance of pathological grade rather than molecular subtypes in guiding chemotherapy decisions for PD-IDC and emphasize the essential role of radiotherapy for patients with PD-IDC at the T4 and N3 stages. Our study offers valuable insights into the prognosis of patients with PD-IDC and contributes to their prognostic prediction, enhancing clinical management by physicians.

Materials and methods

Data source and study design

Figure 1 illustrates the workflow of the study design and analyses conducted in this study. The data on patients with PD-IDC used in this study were sourced from the SEER database (SEER research data, 17 Regs, Nov 2022 Sub [2010-2020]; version 8.4.1), which is publicly accessible. The inclusion criteria were defined as follows: (1) all patients had confirmed evi-

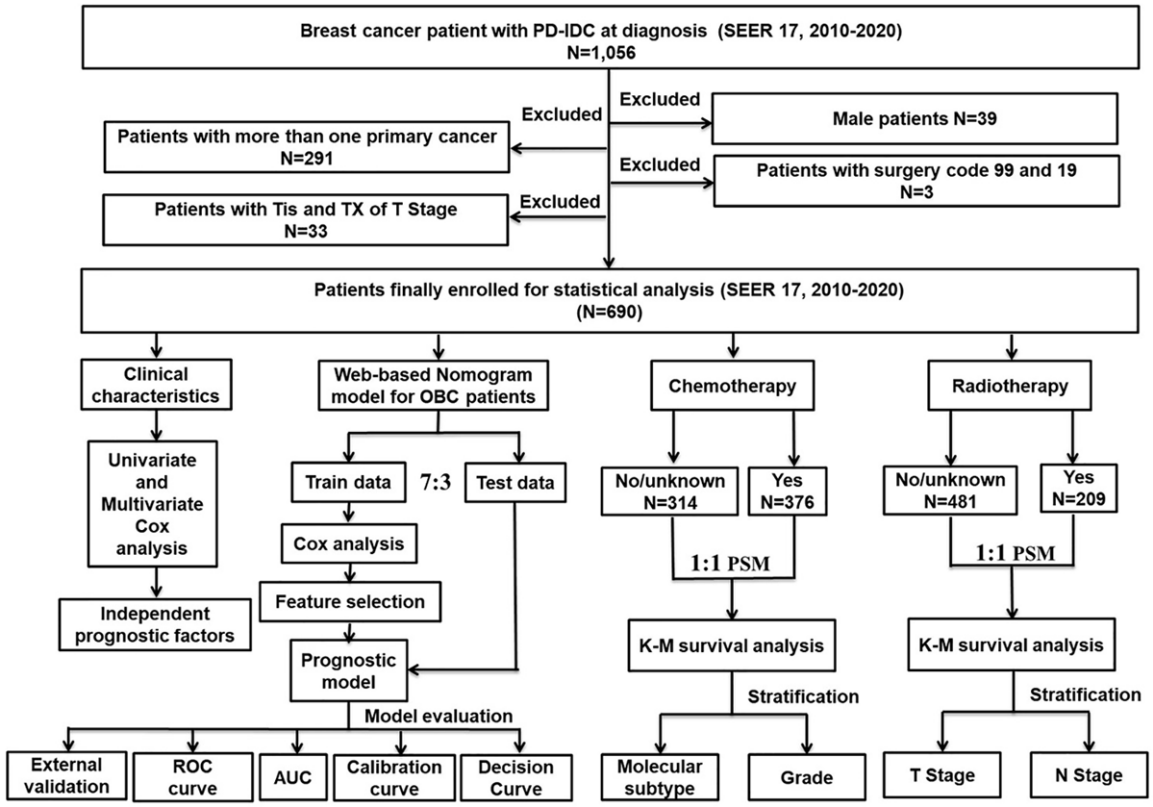


Figure 1. The flowchart detailed the procedure for carrying out the study and analysis of data. PD-IDC, Paget's disease-invasive ductal carcinoma; SEER, Surveillance Epidemiology and End Results; ROC, receiver operating characteristic; AUC, area under the curve.

dence of a morphological and histopathological diagnosis, specifically PD and IDC of the breast (8541/3), as per the International Classification of Cancer Diseases, Third Revision (ICD-O-3); (2) patients were aged ≥ 18 years. The exclusion criteria were defined as follows: (1) male patients; (2) patients with > 1 primary cancer diagnosis; (3) patients with surgery codes 99 and 19, no specimen was sent to pathology for surgical events coded 19 and unknown if surgery performed coded 99; (4) patients with Tis and TX according to the American Joint Committee on Cancer (AJCC); and (5) patients for whom survival time was unknown. Follow-up was conducted until the occurrence of patient death, loss of follow-up, or December 31, 2020.

Nomogram construction and validation

Patients with PD-IDC were randomly divided into two sets: a training set (n = 482) and a testing set (n = 208), following a 7:3 ratio. Multivariable Cox regression analyses were performed on both the entire set and the training set. Characteristics that showed statistical significance in the multivariable Cox regression analyses, such as age, marital status, grade, T stage, N stage, and the presence of bone metastases, were included in our Cox regression models. Subsequently, we utilized the NynNom package and Shiny to construct a web-based nomogram for predicting the OS of patients with PD-IDC. The performance of the nomogram was assessed using various metrics, including fixed-time point area under the curve (AUC) values at three- and five-year intervals, time-dependent AUC, receiver operating characteristic curve (ROC) curves, and calibration curves, to evaluate its discrimination and calibration capabilities. Additionally, decision curve analysis (DCA) was used to assess the clinical utility and benefit of the nomogram.

External validation

For further validation of the web-based nomogram, we gathered data from 12 patients diag-

nosed with PD-IDC between August 2010 and December 2022 from our hospital and sister hospitals. The exclusion criteria for patient selection were consistent with those applied in SEER. We then employed various metrics, including fixed-time point AUC values at three- and five-year intervals, time-dependent AUC, ROC curves, calibration curves, and DCA, to evaluate the performance of the nomogram using our external dataset.

Statistical analysis

Univariate and multivariate Cox regression analyses were used to investigate the risk factors and independent prognostic factors. Among the patients with PD-IDC, those who underwent chemotherapy or radiotherapy and those who did not were matched on a 1:1 basis through PSM, utilizing the variables identified in the univariate Cox regression analyses. This approach helps to mitigate confounding factors, facilitating the comparison of treatment effects across different subgroups [22]. Subsequently, Kaplan-Meier (K-M) curve analysis was conducted [23], stratified by grade, molecular subtype, and T and N stages, within the PSM-adjusted population. All statistical analyses in this study were performed using R software (version 4.1.3), with statistical significance defined as $P < 0.05$.

Ethics statement

Ethical review and approval were waived for this study due to the fact that the data are fully de-identified and no intervention on patients was performed.

Results

Clinical characteristics of patients with PD-IDC

In total, data from 690 patients with PD-IDC were extracted from the SEER database between 2010 and 2020. The clinicopathological features of these patients are summarized in **Table 1** and detailed below. The median age of patients with PD-IDC was 60 years, with a mean age of 59.92±15.16 years. Regarding marital status, 51.45% of the patients were married, while 16.96% were single. In terms of ethnicity, 76.23% of the patients were of white ethnicity, and 11.16% were of black ethnicity. Approximately 23.48% of the patients commenced therapy immediately following diagnosis, while 75.07% began therapy > 1 month after diagnosis. Approximately 39.13% of patients reported a yearly family income exceeding 750,000 USD. Tumor grades were distributed, with 36.23% having grade I/II tumors and 54.06% having grade III/IV tumors. Concerning molecular subtypes, HR-/HER2- constituted 5.51%, followed by HR-/HER2+ (30.43%), HR+/HER2- (28.12%), and HR+/HER2+ (26.96%). Staging classifications showed that 55.51% were at the T1 stage, 23.48% at the T2 stage, 6.81% at the T3 stage, and 14.20% at the T4 stage. Nodal involvement was categorized as 56.81% at stage N0, 27.97% at stage N1, 8.70% at stage N2, and 6.52% at stage N3. For chemotherapy and radiotherapy, 33.33% of patients received chemotherapy only, 9.13% underwent radiotherapy only, and 21.16% of patients received both chemotherapy and radiotherapy. Nearly all patients (96.38%) underwent surgery. The prevalence of metastases was as follows: bone metastases in 3.04% of patients, liver metastases in 1.74%, lung metastases in 2.03%, and no cases of brain metastases.

Univariable and multivariable Cox regression analysis

To identify significant factors influencing OS and BCSS in patients with PD-IDC, we conducted a univariable Cox regression analysis. This analysis encompassed variables such as age at diagnosis, marital status, race, duration from diagnosis to therapy initiation, median family income (adjusted for inflation), tumor grade, molecular subtypes, T and N stages, as well as treatment details and the presence of distant metastases (**Table 2**).

Furthermore, we conducted multivariable Cox regression analyses to eliminate confounding factors and identify the independent factors associated with both OS and BCSS (**Table 2**). The analysis revealed that each additional year of age, high tumor grade, advanced T stages, and the presence of bone metastases were all significantly correlated with worse outcomes for both OS and BCSS. In comparison to marital status, being widowed was associated with poorer OS but did not affect BCSS. Conversely,

Table 1. Baseline characteristics of PD-IDC patients in overall, train and test sets

Characteristic	Level	Overall (n = 690) Cases (%)	Train set (n = 482) Cases (%)	Test set (n = 208) Cases (%)	*P*
Age (mean ± SD)	24~90	59.92±15.16	60.13±15.13	59.44±15.26	0.583
Marriage status	Marital	355 (51.45)	246 (51.04)	109 (52.40)	0.608
	Divorced/Separated	92 (13.33)	59 (12.24)	33 (15.87)	
	Single	117 (16.96)	84 (17.43)	33 (15.87)	
	Unknown	32 (4.64)	23 (4.77)	9 (4.33)	
	Widowed	94 (13.62)	70 (14.52)	24 (11.54)	
Race	White	526 (76.23)	370 (76.76)	156 (75.00)	0.704
	Black	77 (11.16)	55 (11.41)	22 (10.58)	
	Other	83 (12.03)	55 (11.41)	28 (13.46)	
	Unknown	4 (0.58)	2 (0.41)	2 (0.96)	
Months from diagnosis to therapy	0	162 (23.48)	115 (23.86)	47 (22.60)	0.753
	1+	518 (75.07)	361 (74.90)	157 (75.48)	
	Unknown	10 (1.45)	6 (1.24)	4 (1.92)	
Median household income (inflation ajusted)	$44,999-	37 (5.36)	26 (5.39)	11 (5.29)	0.709
	$45,000-$54,999	66 (9.57)	42 (8.71)	24 (11.54)	
	$55,000-$64,999	136 (19.71)	100 (20.75)	36 (17.31)	
	$65,000-$74,999	181 (26.23)	125 (25.93)	56 (26.92)	
	$75,000+	270 (39.13)	189 (39.21)	81 (38.94)	
Grade	Grade I/II	250 (36.23)	175 (36.31)	75 (36.06)	0.935
	Grade III/Grade IV	373 (54.06)	259 (53.73)	114 (54.81)	
	Unknown	67 (9.71)	48 (9.96)	19 (9.13)	
Subtypes	HR-/HER2-	38 (5.51)	29 (6.02)	9 (4.33)	0.623
	HR-/HER2+	210 (30.43)	153 (31.74)	57 (27.40)	
	HR+/HER2-	194 (28.12)	130 (26.97)	64 (30.77)	
	HR+/HER2+	186 (26.96)	128 (26.56)	58 (27.88)	
	Unknown	62 (8.99)	42 (8.71)	20 (9.62)	
T Stage	T1	383 (55.51)	269 (55.81)	114 (54.81)	0.331
	T2	162 (23.48)	112 (23.24)	50 (24.04)	
	T3	47 (6.81)	28 (5.81)	19 (9.13)	
	T4	98 (14.20)	73 (15.15)	25 (12.02)	
N Stage	N0	392 (56.81)	266 (55.19)	126 (60.58)	0.262
	N1	193 (27.97)	145 (30.08)	48 (23.08)	
	N2	60 (8.70)	39 (8.09)	21 (10.10)	
	N3	45 (6.52)	32 (6.64)	13 (6.25)	
Treatment combination (Radiation or Chemotherapy)	None/Unknown	251 (36.38)	168 (34.85)	83 (39.90)	0.536
	Radiation only	63 (9.13)	46 (9.54)	17 (8.17)	
	Chemotherapy only	230 (33.33)	167 (34.65)	63 (30.29)	
	Radiation and Chemotherapy	146 (21.16)	101 (20.95)	45 (21.63)	
Surgery	No	25 (3.62)	17 (3.53)	8 (3.85)	1
	Yes	665 (96.38)	465 (96.47)	200 (96.15)	
Bone metastasis	No	666 (96.52)	468 (97.10)	198 (95.19)	0.275
	Yes	21 (3.04)	13 (2.70)	8 (3.85)	
	Unknown	3 (0.43)	1 (0.21)	2 (0.96)	
Liver metastasis	No	674 (97.68)	472 (97.93)	202 (97.12)	0.665
	Yes	12 (1.74)	8 (1.66)	4 (1.92)	
	Unknown	4 (0.58)	2 (0.41)	2 (0.96)	
Lung metastasis	No	670 (97.10)	471 (97.72)	199 (95.67)	0.131
	Yes	14 (2.03)	9 (1.87)	5 (2.40)	
	Unknown	6 (0.87)	2 (0.41)	4 (1.92)	
Brain metastasis	No	686 (99.42)	480 (99.59)	206 (99.04)	0.748
	Yes	0 (0)	0 (0)	0 (0)	
	Unknown	4 (0.58)	2 (0.41)	2 (0.96)	

Table 2. Univariate and multivariate Cox analysis of PD-IDC characteristics

	Univariate Cox analysis						Multivariate Cox analysis					
	OS			BCSS			OS			BCSS		
	HR	95% CI	*P* Value	HR	95% CI	*P* Value	HR	95% CI	*P* Value	HR	95% CI	*P* Value
Age												
24~90	1.06	1.05-1.07	***	1.03	1.01-1.04	***	1.05	1.03-1.07	***	1.04	1.02-1.06	***
Marriage status												
Married	reference			reference			reference			reference		
Divorced/Separated	1.29	0.73-2.27	0.38	0.81	0.36-1.83	0.61	1.67	0.88-3.17	0.12	1.07	0.46-2.51	0.87
Single	2.25	1.43-3.52	***	2.12	1.23-3.65	**	1.66	0.94-2.94	0.08	1.32	0.70-2.48	0.39
Widowed	4.91	3.30-7.30	***	1.93	1.03-3.60	*	2.30	1.33-3.97	**	1.05	0.47-2.31	0.91
Race												
White	reference			reference			/	/	/	/	/	/
Black	1.08	0.66-11.77	0.77	1.26	0.66-2.40	0.48	/	/	/	/	/	/
Other	0.66	0.36-1.19	0.17	0.78	0.36-1.71	0.53	/	/	/	/	/	/
Months from diagnosis to therapy												
0 month	reference			reference			/	/	/	/	/	/
≥ 1 month	0.78	0.55-1.10	0.16	0.72	0.45-1.15	0.17	/	/	/	/	/	/
Median household income (inflation ajusted)												
< 45,000$	reference			reference			reference			reference		
45,000-54,999$	0.34	0.16-0.68	**	0.38	0.14-1.02	0.06	0.48	0.20-1.19	0.11	0.48	0.168-1.41	0.18
55,000-64,999$	0.54	0.31-0.94	*	0.61	0.28-1.33	0.21	0.71	0.32-1.54	0.38	0.67	0.29-1.53	0.34
65,000-74,999$	0.40	0.23-0.70	**	0.46	0.21-0.99	*	0.70	0.33-1.50	0.36	0.72	0.31-1.65	0.44
> 74,999$	0.28	0.16-0.50	***	0.30	0.14-0.67	**	0.50	0.23-1.09	0.08	0.38	0.16-0.89	*
Grade												
Grade I/Grade II	reference			reference						reference		
Grade III/Grade IV	1.74	1.21-2.50	**	2.15	1.28-3.63	**	1.97	1.21-3.19	**	2.11	1.16-3.84	*
Subtypes												
HR-/HER2-	reference			reference			reference			/	/	/
HR-/HER2+	0.39	0.21-0.76	**	0.58	0.21-1.59	0.29	0.61	0.26-1.45	0.26	/	/	/
HR+/HER2-	0.65	0.35-1.22	0.18	1.00	0.38-2.61	0.99	0.78	0.33-1.85	0.57	/	/	/
HR+/HER2+	0.80	0.43-1.47	0.47	1.12	0.43-2.92	0.82	1.16	0.52-2.57	0.72	/	/	/
T Stage												
T1	reference			reference			reference			reference		
T2	2.06	1.36-3.11	***	5.50	2.71-11.17	***	1.61	0.93-2.77	0.09	3.07	1.37-6.88	**
T3	2.63	1.48-4.69	**	6.94	2.88-16.75	***	4.20	1.83-9.62	***	5.44	1.88-15.71	**
T4	4.84	3.24-7.23	***	15.28	7.75-30.10	***	3.00	1.63-5.53	***	5.96	2.55-13.90	***

N Stage												
N0	reference			reference			reference			reference		
N1	1.61	1.10-2.36	*	3.44	1.91-6.19	***	1.49	0.90-2.47	0.12	1.73	0.87-3.44	0.12
N2	2.42	1.50-3.90	***	5.46	2.75-10.84	***	1.80	0.92-3.50	0.09	1.96	0.88-4.37	0.10
N3	2.95	1.80-4.84	***	8.77	4.56-16.86	***	2.14	0.97-4.64	0.060	2.42	1.06-5.54	*
Treatment combination (Radiation or Chemotherapy)												
No/Unknown	reference			reference			reference			reference		
Radiation only	055	0.29-1.04	0.068	0.66	0.25-1.70	0.39	0.81	0.38-1.76	0.60	/	/	/
Chemotherapy only	0.46	0.31-0.67	***	0.79	0.46-1.30	0.37	0.64	0.35-1.16	0.14	/	/	/
Radiation and Chemotherapy	0.44	0.28-0.69	***	0.96	0.54-1.70	0.89	0.54	0.28-1.07	0.078	/	/	/
Surgery												
No	reference			reference			reference			reference		
Yes	0.28	0.15-0.54	***	0.19	0.09-0.39	***	0.64	0.21-1.95	0.43	0.79	0.24-2.62	0.70
Bone metastasis												
No	reference						reference					
Yes	7.22	4.21-12.37	***	12.92	7.21-23.16	***	5.03	1.88-13.48	**	3.94	1.59-9.75	**
Liver metastasis												
No	reference			reference			reference			reference		
Yes	8.27	4.33-15.79	***	14.62	7.49-28.54	***	1.77	0.45-7.08	0.41	2.31	0.70-7.57	0.17
Lung metastasis												
No	reference			reference			reference			reference		
Yes	6.30	3.31-12.00	***	13.00	6.66-25.35	***	1.80	0.62-5.23	0.28	2.61	1.04-6.55	*

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

Table 3. Performance of different prognostic models (AUC of the ROC curve)

Model parameter	Train set		Test set	
	3-year survival	5-year survival	3-year survival	5-year survival
Age+Marital+Grade+T+N+bone	0.852	0.853	0.831	0.841
Age+Marital+Grade+T+bone	0.837	0.837	0.826	0.841
Age+Marital+Grade+T+N	0.850	0.851	0.820	0.822
Age+Marital+Grade+T	0.843	0.835	0.803	0.823

ROC, receiver operating characteristic; AUC, area under the curve.

a higher family income (> 74,999 *vs.* < 45,000 USD) was associated with improved BCSS but did not impact OS in patients with PD-IDC. Only for BCSS, patients at the N3 stage showed a more unfavorable prognosis than those at the N0 stage, while the N1 and N2 stages exhibited similar outcomes to the N0 stage. Surprisingly, the molecular subtype, surgical intervention, chemotherapy, radiotherapy and even combination of chemotherapy and radiotherapy did not emerge as independent factors influencing the prognosis of patients with PD-IDC.

Web-based nomogram construction and assessment

Given the aforementioned considerations, patients were randomly divided into training and testing datasets in a 7:3 ratio (**Table 1**). Within the training dataset, we conducted both univariate and multivariate Cox regression analyses once again (Supplementary Table 1). By combining the results of the multivariable Cox regression analyses from the entire dataset and the training set, we selected six independent prognostic factors to serve as the key model features. Subsequently, we examined the AUC values for three- and five-year intervals for various combinations and determined that age at diagnosis, marital status, tumor grade, T and N stage, and the presence of bone metastasis were the most effective model features (**Table 3**) for predicting the OS of patients with PD-IDC at both three and five years. Consequently, we constructed a web-based nomogram (**Figure 2A-D**, available at https://quxjtu.shinyapps.io/PD-IDCapp/), which facilitates the drawing of survival curves (**Figure 2**), calculating 95% confidence intervals (CIs) for survival probability (**Figure 2B**), presenting a summary of each prediction (**Figure 2C**), and reviewing the model features (**Figure 2D**). Importantly, this web-based nomogram offers the capability to predict OS for patients with PD-IDC at multiple time points.

Subsequent to the construction of the nomogram, we conducted an assessment of its performance. Predicted ROC curves were generated, and their corresponding AUC values were calculated at specific time points (3 and 5 years) for both the training (**Figure 3A**) and testing datasets (**Figure 3B**). Our web-based nomogram exhibited extraordinarily high discriminative ability in predicting the survival of patients with PD-IDC, achieving AUCs of 0.852 and 0.831 for the three-year prediction in the training and testing sets, respectively. Similarly, for the five-year prediction, the AUC values were 0.853 and 0.841 for the training and testing sets, respectively (**Figure 3A** and **3B**). The time-dependent AUCs illustrated in both the training (**Figure 3C**) and testing sets (**Figure 3D**) further demonstrated the model's stability over time while maintaining excellent discriminative power. To further evaluate the accuracy of our nomogram, calibration curves were plotted [24], indicating a remarkable alignment between the predicted and observed values for both the training and testing sets (**Figure 4A-D**). Additionally, we analyzed the clinical applicability of our model through DCA [25]. The DCA revealed a substantial threshold probability range and favorable net benefit in the decision curves for predicting three- and five-year OS rates among patients with PD-IDC in both the training and testing sets (**Figure 5A-D**). In summary, our web-based nomogram demonstrated outstanding performance across various assessment measures.

External validation for the nomogram

To further validate the efficacy of our nomogram, we collected relevant information from an additional 12 patients with PD-IDC from both our hospital and sister hospitals (Supplementary Table 2). The results indicated that our nomogram maintained strong performance in this external dataset, achieving AUC values of 0.892 and 0.914 for the three- and five-year predictions, respectively (**Figure 6A**). Furthermore, the time-dependent AUC (**Figure**

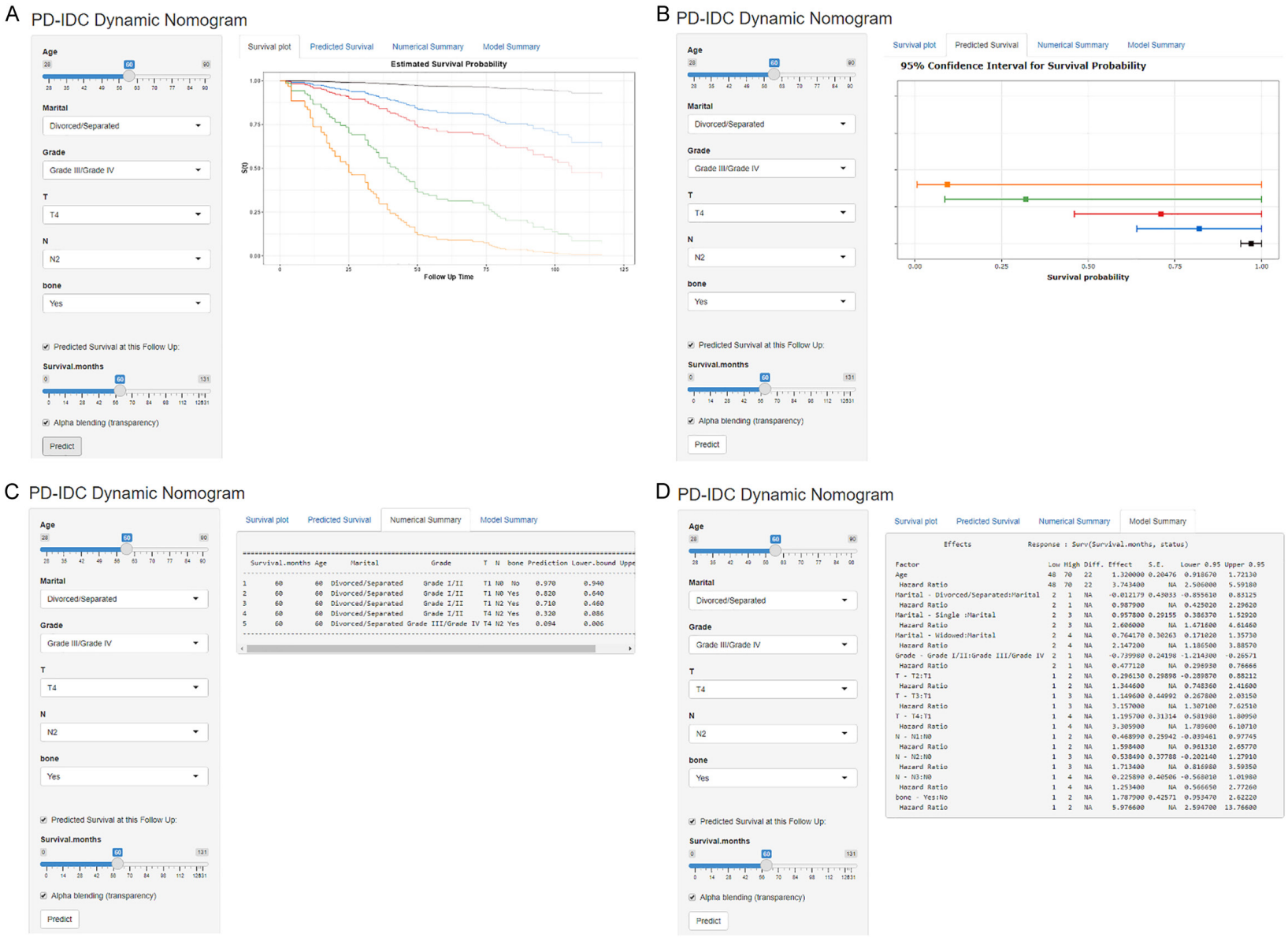


Figure 2. Web-based prognostic nomogram for patients with PD-IDC (available at: https://quxjtu.shinyapps.io/PD-IDCapp/). A. Enables the creation of survival curves; B. Calculates 95% confidence intervals for survival probability; C. Provides a summary of each prediction; D. Allows for easy review of model features. PD-IDC, Paget's disease-invasive ductal carcinoma.

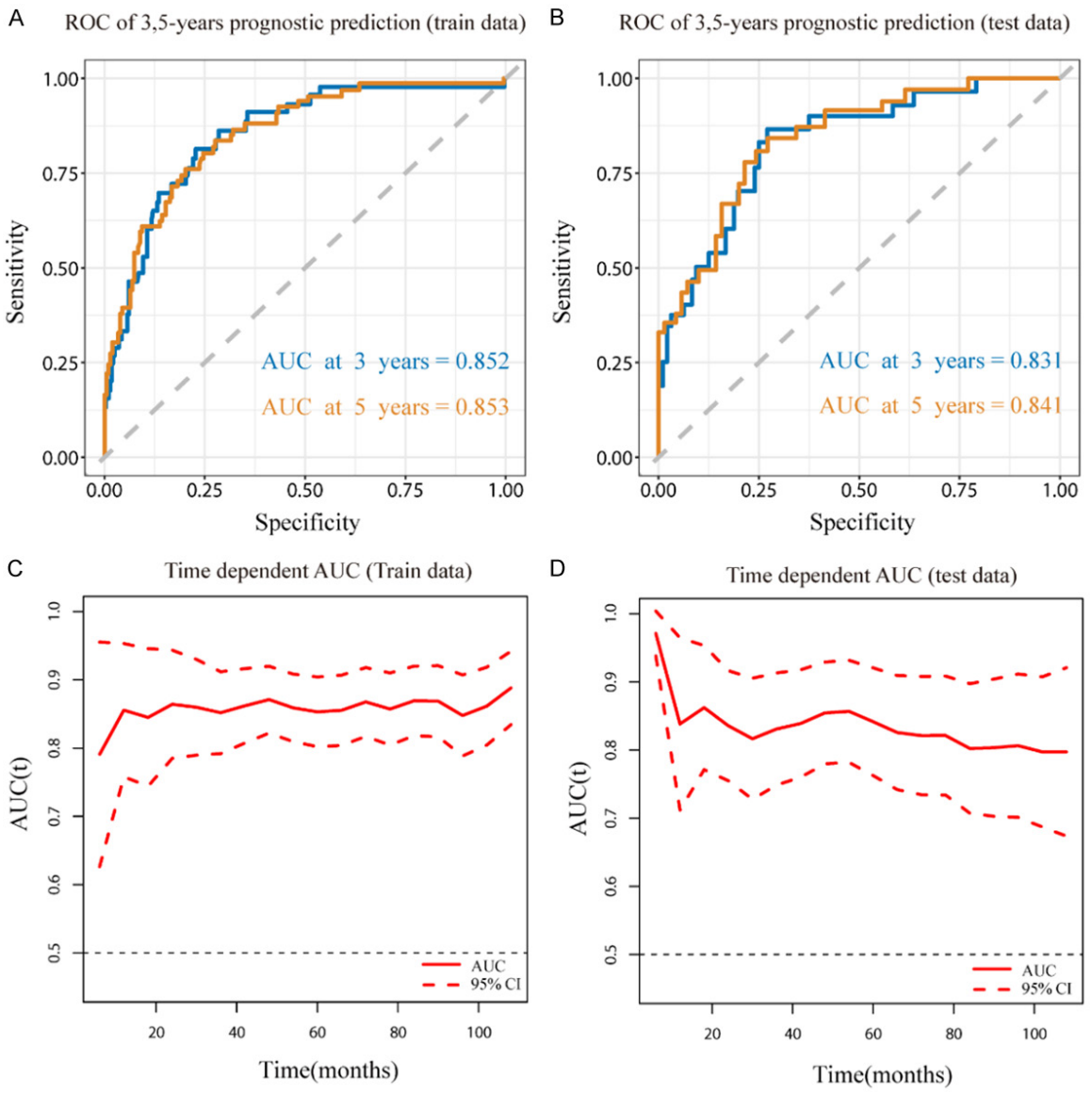


Figure 3. ROC curve analysis for the web-based prognostic nomogram. ROC curve illustrating the performance of the prognostic model: (A) Three- and five-year predictions in the training dataset; (B) Three- and five-year predictions in the testing dataset; (C) Time-dependent AUC in the training dataset; (D) Time-dependent AUC in the testing dataset. ROC, receiver operating characteristic; AUC, area under the curve.

6B), calibration curve results (**Figure 6C** and **6D**), and DCA outcomes (**Figure 6E** and **6F**) all confirmed the robust discrimination, calibration capacity, and clinical applicability of the nomogram within this external dataset.

Benefit of chemotherapy in patients with PD-IDC by grade and molecular subtype

Precipitously, in our multivariable Cox regression analysis (**Table 2**), chemotherapy did not emerge as an independent prognostic factor for patients with PD-IDC. Consequently, we performed additional stratification analyses to investigate which factors might influence the efficacy of chemotherapy. A comparison of baseline characteristics between patients who underwent chemotherapy and those who did not reveal noticeable differences (**Table 4**). Therefore, we used PSM to balance their baseline characteristics (**Table 4**).

According to the PSM-adjusted data, there is no significant difference in both OS (**Figure 7A**)

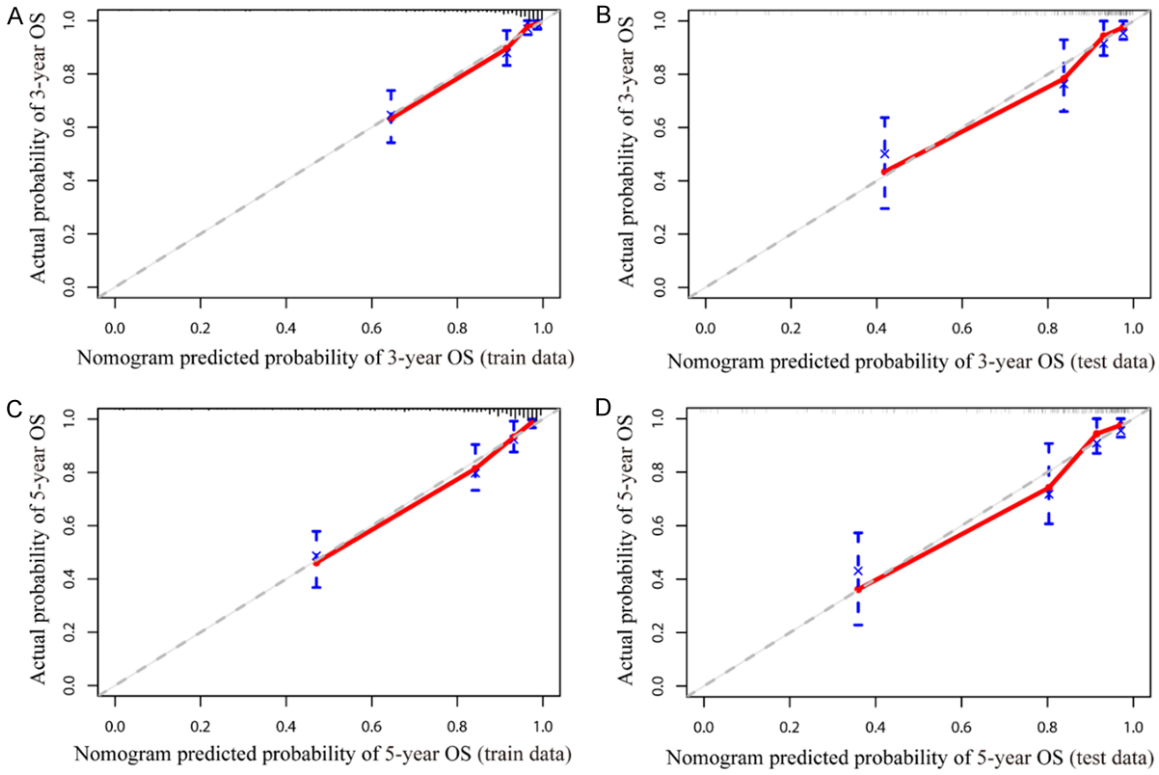


Figure 4. Calibration curve assessment for the web-based prognostic nomogram. Calibration curves evaluating the web-based prognostic nomogram's predictive accuracy for: (A) Three-year OS in the training dataset; (B) Three-year OS in the testing dataset; (C) Five-year OS in the training dataset; (D) Five-year OS in the testing dataset. OS, overall survival.

and BCSS (**Figure 7B**) between patients with PD-IDC who underwent chemotherapy and those who did not. Typically, different molecular types exhibit varying sensitivities to chemotherapy. However, when we conducted stratification analysis by molecular subtype (Supplementary Figure 1A-H), it became evident that chemotherapy failed to improve the OS of patients with PD-IDC across all molecular subtypes (Supplementary Figure 1A-D). Although patients with HR-/Her-2+ showed improved BCSS after chemotherapy (Supplementary Figure 1F), further analysis using univariable Cox regression in all patients with HR-Her-2+ showed that chemotherapy was not a prognostic factor for BCSS (P = 0.0603, hazard ratio [HR] = 0.3898, 95% CI = 0.1459-1.042). Hence, it is apparent that the molecular subtype is not an effective predictor of chemotherapy efficacy for both OS and BCSS.

On the other hand, the pathological grade can sometimes offer a preliminary assessment of chemotherapy sensitivity [26]. Consequently, we conducted a stratification analysis by grade, revealing that chemotherapy could improve the OS and BCSS of patients with PD-IDC only in the grade III/IV group (**Figure 8B** and **8D**) but not in the grade I/II group (**Figure 8A** and **8C**). All the data suggests that it is the pathological grade, rather than the molecular subtype, that determines the efficacy of chemotherapy in improving the prognosis for patients with PD-IDC.

Benefit of radiotherapy in patients with PD-IDC by T and N stage

Considering the previous findings, we further investigated the factors that might influence the efficacy of radiotherapy for patients with PD-IDC. Employing the same PSM method, we successfully balanced the baseline characteristics between patients who received radiotherapy and those who did not (**Table 5**).

Based on the PSM-adjusted data, there were no discernible differences in either OS (**Figure 9A**) or BCSS (**Figure 9B**) between patients with PD-IDC who underwent radiotherapy and th-

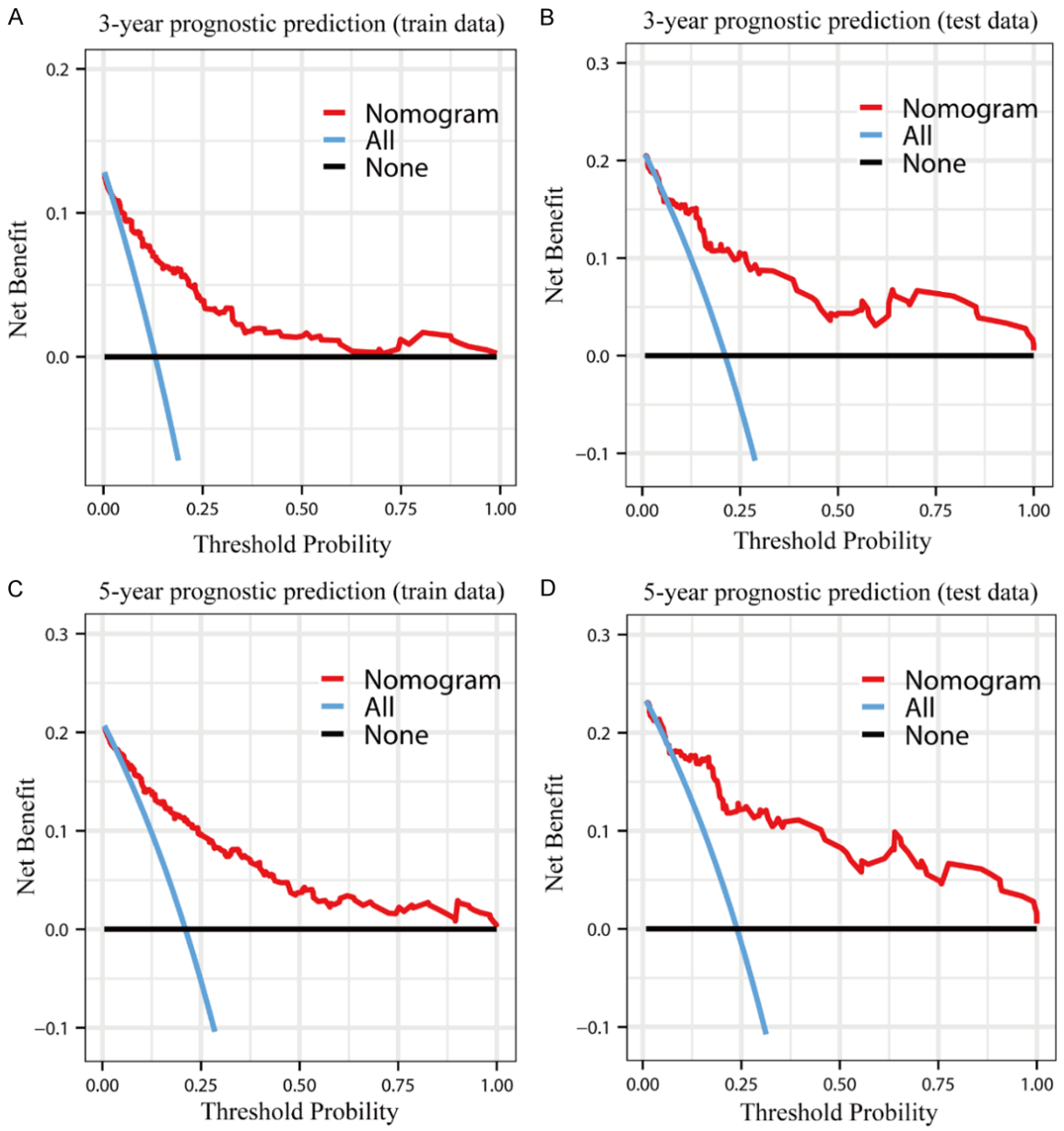


Figure 5. Decision curve analysis of the web-based prognostic nomogram. Decision curves demonstrating the clinical utility of the web-based prognostic nomogram in predicting: (A) Three-year OS in the training dataset; (B) Three-year OS in the testing dataset; (C) Five-year OS in the training dataset; (D) Five-year OS in the testing dataset. OS, overall survival.

ose who did not. Further stratified K-M survival analysis showed that the benefits of radiotherapy in terms of both OS and BCSS (**Figures 10A-H** and **11A-H**) were confined to only those patients with PD-IDC with stage T4 (**Figure 10D** and **10H**) and N3 (**Figure 11D** and **11H**). No other characteristics were found to influence the impact of radiotherapy on prognosis (data not shown).

The combination of chemotherapy and radiotherapy can not improve the survival of patients with PD-IDC

We have found that compared with the patients without any chemotherapy and radiotherapy, the combination of chemotherapy and radiotherapy was not an independent prognostic factor (**Table 2**), but what about compared with

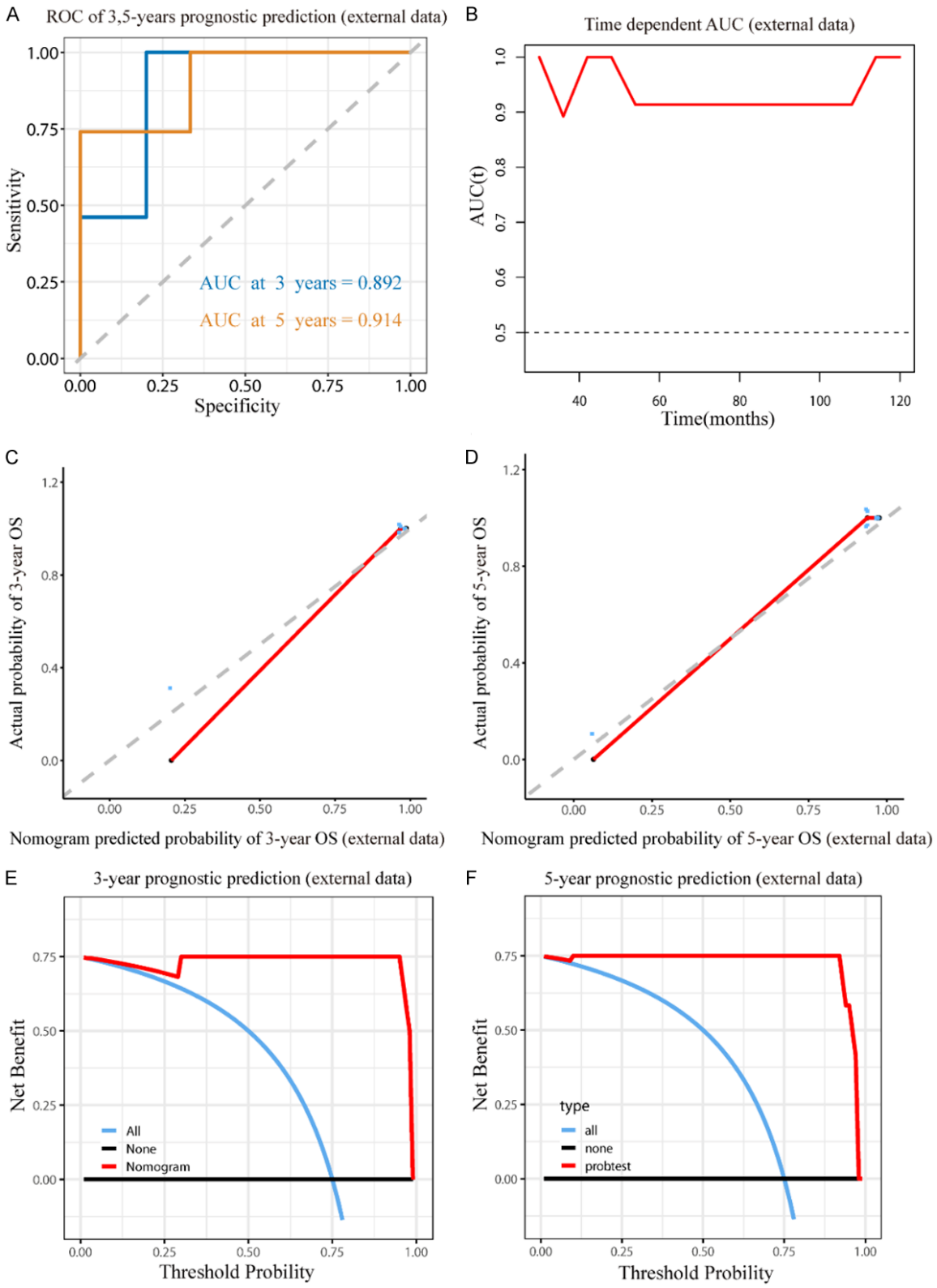


Figure 6. External validation of the web-based prognostic nomogram. Validation of the prognostic model's performance in external data, including: (A) ROC curve for three-, and five-year predictions in the external dataset; (B) Time-dependent AUC in the external dataset; (C) Calibration curve for three-year predictions in the external dataset; (D) Calibration curve for five-year predictions in the external dataset; (E) Decision curves for three-year predictions in the external dataset; (F) Decision curves for five-year predictions in the external dataset. ROC, receiver operating characteristic; AUC, area under the curve.

Table 4. Comparison of patient features by chemotherapy before and after propensity score matching (PSM)

Characteristics	Unmatched Cohort			1:1 propensity score matched (PSM) Cohort		
	Chemotherapy not given	Chemotherapy	Unadjusted	Chemotherapy not given	Chemotherapy	PSM-adjusted
	N = 314 (%)	N = 376 (%)	*P* value	N = 191 (%)	N = 191 (%)	*P* value
Age			***			0.252
< 50	42 (13.38)	148 (39.36)		41 (21.47)	58 (30.37)	
50-59	55 (17.52)	97 (25.80)		45 (23.56)	48 (25.13)	
60-69	79 (25.16)	81 (21.54)		57 (29.84)	49 (25.65)	
70-79	61 (19.43)	43 (11.44)		38 (19.90)	29 (15.18)	
80+	77 (24.52)	7 (1.86)		10 (5.24)	7 (3.66)	
Marriage status			***			0.9
Married	142 (45.22)	213 (56.65)		98 (51.31)	105 (54.97)	
Divorced/Separated	43 (13.69)	49 (13.03)		30 (15.71)	26 (13.61)	
Single	48 (15.29)	69 (18.35)		32 (16.75)	34 (17.80)	
Widowed	67 (21.34)	27 (7.18)		21 (10.99)	17 (8.90)	
Unknown	14 (4.46)	18 (4.79)		10 (5.24)	9 (4.71)	
Race			**			0.343
White	254 (80.89)	272 (72.34)		150 (78.53)	139 (72.77)	
Black	22 (7.01)	55 (14.63)		15 (7.85)	25 (13.09)	
Other	35 (11.15)	48 (12.77)		24 (12.57)	26 (13.61)	
Unknown	3 (0.96)	1 (0.27)		2 (1.05)	1 (0.52)	
Months from diagnosis to therapy			0.081			0.299
0 month	82 (26.11)	80 (21.28)		43 (22.51)	39 (20.42)	
≥ 1 month	225 (71.66)	293 (77.93)		142 (74.35)	150 (78.53)	
Unknown	7 (2.23)	3 (0.80)		6 (3.14)	2 (1.05)	
Median household income (inflation ajusted)			0.079			
$44,999-	11 (3.50)	26 (6.91)		6 (3.14)	10 (5.24)	0.872
$45,000-$54,999	26 (8.28)	40 (10.64)		19 (9.95)	17 (8.90)	
$55,000-$64,999	65 (20.70)	71 (18.88)		32 (16.75)	31 (16.23)	
$65,000-$74,999	94 (29.94)	87 (23.14)		54 (28.27)	51 (26.70)	
$75,000+	118 (37.58)	152 (40.43)		80 (41.88)	82 (42.93)	
Grade						0.745
Grade I/II	129 (41.08)	121 (32.18)	***	75 (39.27)	68 (35.60)	
Grade III/Grade IV	145 (46.18)	228 (60.64)		101 (52.88)	106 (55.50)	
Unknown	40 (12.74)	27 (7.18)		15 (7.85)	17 (8.90)	
Subtypes			***			
HR-/HER2-	16 (5.10)	22 (5.85)		10 (5.24)	12 (6.28)	0.925
HR-/HER2+	71 (22.61)	139 (36.97)		59 (30.89)	60 (31.41)	
HR+/HER2-	105 (33.44)	89 (23.67)		64 (33.51)	57 (29.84)	
HR+/HER2+	76 (24.20)	110 (29.26)		45 (23.56)	50 (26.18)	
Unknown	46 (14.65)	16 (4.26)		13 (6.81)	12 (6.28)	
T Stage			***			
T1	209 (66.56)	174 (46.28)		122 (63.87)	109 (57.07)	0.578
T2	60 (19.11)	102 (27.13)		38 (19.90)	47 (24.61)	
T3	11 (3.50)	36 (9.57)		9 (4.71)	11 (5.76)	
T4	34 (10.83)	64 (17.02)		22 (11.52)	24 (12.57)	

N Stage			***			0.059
N0	230 (73.25)	162 (43.09)		124 (64.92)	105 (54.97)	
N1	64 (20.38)	129 (34.31)		52 (27.23)	56 (29.32)	
N2	11 (3.50)	49 (13.03)		9 (4.71)	22 (11.52)	
N3	9 (2.87)	36 (9.57)		6 (3.14)	8 (4.19)	
Radiation			***			0.301
No/Unknown	251 (79.94)	230 (61.17)		144 (75.39)	134 (70.16)	
Yes	63 (20.06)	146 (38.83)		47 (24.61)	57 (29.84)	
Surgery			0.96			0.785
No	12 (3.82)	13 (3.46)		8 (4.19)	6 (3.14)	
Yes	302 (96.18)	363 (96.54)		183 (95.81)	185 (96.86)	
Bone metastases			0.403			0.966
No	305 (97.13)	361 (96.01)		183 (95.81)	182 (95.29)	
Yes	7 (2.23)	14 (3.72)		7 (3.66)	8 (4.19)	
Unknown	2 (0.64)	1 (0.27)		1 (0.52)	1 (0.52)	
Liver metastases			0.649			0.904
No	306 (97.45)	362 (96.28)		184 (96.34)	185 (96.86)	
Yes	4 (1.27)	8 (2.13)		4 (2.09)	4 (2.09)	
Unknown	4 (1.27)	6 (1.60)		3 (1.57)	2 (1.05)	
Lung metastases			0.569			0.356
No	304 (96.82)	366 (97.34)		182 (95.29)	187 (97.91)	
Yes	6 (1.91)	8 (2.13)		6 (3.14)	3 (1.57)	
Unknown	4 (1.27)	2 (0.53)		3 (1.57)	1 (0.52)	
Brain metastases			0.494			1
No	311 (99.04)	375 (99.73)		189 (98.95)	190 (99.48)	
Yes	0 (0.00)	0 (0.00)		0 (0.00)	0 (0.00)	
Unknown	3 (0.96)	1 (0.27)		2 (1.05)	1 (0.52)	

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

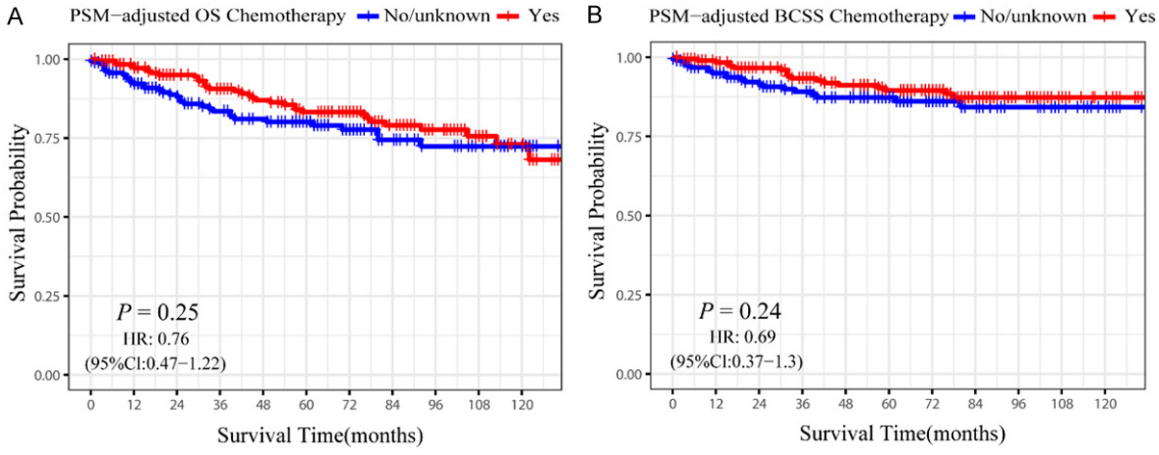


Figure 7. OS and BCSS analysis of patients with PD-IDC undergoing chemotherapy after PSM adjustment. PSM-adjusted Kaplan-Meier (K-M) survival analysis: (A) OS of patients with PD-IDC who underwent chemotherapy versus those who did not; (B) BCSS of patients with PD-IDC who underwent chemotherapy versus those who did not. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching.

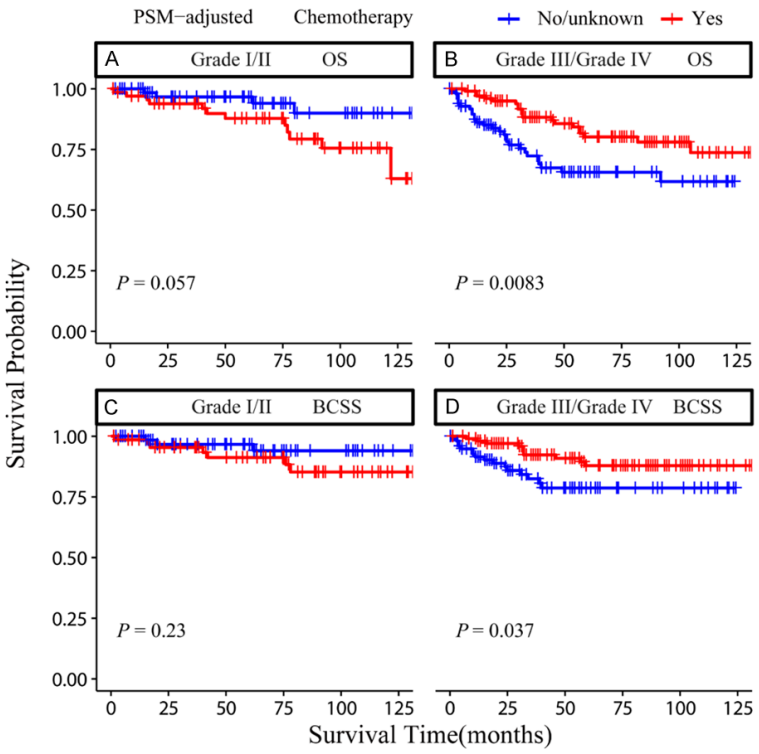


Figure 8. OS and BCSS of patients with PD-IDC who underwent chemotherapy after PSM adjustment (Stratified by grade). A. OS of patients with PD-IDC with grade I/II tumors; B. OS of patients with PD-IDC with grade III/IV tumors; C. BCSS of patients with PD-IDC with grade I/II tumors; D. BCSS of patients with PD-IDC with grade III/IV tumors. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching.

the patients with chemotherapy only or radiotherapy only? Thus, after PSM, we further compared the effect of radiotherapy in chemotherapy group and no-chemotherapy group (**Figure 12A-D**). We also compared the effect of chemotherapy in radiotherapy group and no-radiotherapy group (**Figure 13A-D**). The results showed that patients received both chemotherapy and radiotherapy could not improve the survival, and the curative effect of two treatments do not affect each other.

Discussion

PD of the breast represents an uncommon clinical entity, with approximately 50% to 60% of these patients concurrently diagnosed with IDC [3, 7]. Among the different subtypes of PD, PD-IDC carries the worst prognosis [10]. In comparison to typical IDC, patients with PD-IDC are more likely to have axillary lymph node metastasis [7], which usually results in a poorer prognosis [4, 8-10]. Given its specificity, prognostic prediction becomes particularly crucial for these patients. Unfortunately, previous conventional prognostic models for IDC [13-15] failed to consider the differences between PD-IDC and IDC. Consequently, PD-IDC was excluded from these models, leading to inaccurate assessments of patients with PD-IDC. To address this gap, we constructed a novel web-based nomogram for predicting the prognosis of patients with PD-IDC. Our predictive model is not only highly accurate but also easier for doctors and patients to use than a traditional nomogram. Traditional nomograms [27-30] require manual measurement with a ruler and tedious calculations, whereas our web-based nomogram streamlines the process by simply inputting the required information on the website. Furthermore, our model offers the flexibility to predict prognosis at various time points and provides more comprehensive results, whereas traditional nomograms can only predict outcomes at fixed time points, yielding limited information. Additionally, our study benefits from the most recent SEER update for the year 2020, ensuring that it represents the most up-to-date investigation into the clinical features and prognosis of patients with PD-IDC using SEER data.

Our study successfully identified the independent prognostic factors of patients with PD-IDC, including age, marital status, tumor grade, T and N stages, and the presence of bone metastases. Notably, each additional year of age was significantly associated with poorer OS and BCSS, aligning with a previous study [8]. However, it is worth mentioning that the aforementioned previous study did not explore BCSS. Conversely, another study suggested that patients < 50 years were more likely to experience worse OS [10]. In contrast to marital status, our findings indicated that widowed sta-

Table 5. Comparison of patient features by radiotherapy before and after propensity score matching (PSM)

Characteristics	Unmatched Cohort			1:1 propensity score matched (PSM) Cohort		
	Radiotherapy not given	Radiotherapy	Unadjusted	Radiotherapy not given	Radiotherapy	PSM-adjusted
	N = 481 (%)	N = 209 (%)	*P* value	N = 209 (%)	N = 209 (%)	*P* value
Age			**			0.944
< 50	126 (26.20)	64 (30.62)		59 (28.23)	64 (30.62)	
50-59	108 (22.45)	44 (21.05)		43 (20.57)	44 (21.05)	
60-69	104 (21.62)	56 (26.79)		60 (28.71)	56 (26.79)	
70-79	69 (14.35)	35 (16.75)		34 (16.27)	35 (16.75)	
80+	74 (15.38)	10 (4.78)		13 (6.22)	10 (4.78)	
Marriage status			0.343			0.933
Married	241 (50.10)	114 (54.55)		107 (51.20)	114 (54.55)	
Divorced/Separated	62 (12.89)	30 (14.35)		30 (14.35)	30 (14.35)	
Single	81 (16.84)	36 (17.22)		43 (20.57)	36 (17.22)	
Widowed	74 (15.38)	20 (9.57)		20 (9.57)	20 (9.57)	
Unknown	23 (4.78)	9 (4.31)		9 (4.31)	9 (4.31)	
Race			0.462			0.411
White	362 (75.26)	164 (78.47)		154 (73.68)	164 (78.47)	
Black	52 (10.81)	25 (11.96)		28 (13.40)	25 (11.96)	
Other	64 (13.31)	19 (9.09)		27 (12.92)	19 (9.09)	
Unknown	3 (0.62)	1 (0.48)		0 (0.00)	1 (0.48)	
Months from diagnosis to therapy			*			0.084
0 month	104 (21.62)	58 (27.75)		45 (21.53)	58 (27.75)	
≥ 1 month	367 (76.30)	151 (72.25)		161 (77.03)	151 (72.25)	
Unknown	10 (2.08)	0 (0.00)		3 (1.44)	0 (0.00)	
Median household income (inflation ajusted)			0.71			0.956
$44,999-	25 (5.20)	12 (5.74)		11 (5.26)	12 (5.74)	
$45,000-$54,999	46 (9.56)	20 (9.57)		19 (9.09)	20 (9.57)	
$55,000-$64,999	100 (20.79)	36 (17.22)		38 (18.18)	36 (17.22)	
$65,000-$74,999	129 (26.82)	52 (24.88)		58 (27.75)	52 (24.88)	
$75,000+	181 (37.63)	89 (42.58)		83 (39.71)	89 (42.58)	
Grade			0.089			0.954
Grade I/II	163 (33.89)	87 (41.63)		90 (43.06)	87 (41.63)	
Grade III/Grade IV	266 (55.30)	107 (51.20)		104 (49.76)	107 (51.20)	
Unknown	52 (10.81)	15 (7.18)		15 (7.18)	15 (7.18)	
Subtypes			*			0.271
HR-/HER2-	26 (5.41)	12 (5.74)		13 (6.22)	12 (5.74)	
HR-/HER2+	152 (31.60)	58 (27.75)		72 (34.45)	58 (27.75)	
HR+/HER2-	119 (24.74)	75 (35.89)		58 (27.75)	75 (35.89)	
HR+/HER2+	133 (27.65)	53 (25.36)		49 (23.44)	53 (25.36)	
Unknown	51 (10.60)	11 (5.26)		17 (8.13)	11 (5.26)	
T Stage			***			0.261
T1	292 (60.71)	91 (43.54)		111 (53.11)	91 (43.54)	
T2	105 (21.83)	57 (27.27)		48 (22.97)	57 (27.27)	
T3	24 (4.99)	23 (11.00)		17 (8.13)	23 (11.00)	
T4	60 (12.47)	38 (18.18)		33 (15.79)	38 (18.18)	

N Stage			***			0.078
N0	316 (65.70)	76 (36.36)		93 (44.50)	76 (36.36)	
N1	120 (24.95)	73 (34.93)		78 (37.32)	73 (34.93)	
N2	25 (5.20)	35 (16.75)		22 (10.53)	35 (16.75)	
N3	20 (4.16)	25 (11.96)		16 (7.66)	25 (11.96)	
Chemotherapy			***			0.297
No/Unknown	251 (52.18)	63 (30.14)		74 (35.41)	63 (30.14)	
Yes	230 (47.82)	146 (69.86)		135 (64.59)	146 (69.86)	
Surgery			**			1
No	24 (4.99)	1 (0.48)		0 (0.00)	1 (0.48)	
Yes	457 (95.01)	208 (99.52)		209 (100.00)	208 (99.52)	
Bone metastases			0.511			0.749
No	463 (96.26)	203 (97.13)		205 (98.09)	203 (97.13)	
Yes	15 (3.12)	6 (2.87)		4 (1.91)	6 (2.87)	
Unknown	3 (0.62)	0 (0.00)		0 (0.00)	0 (0.00)	
Liver metastases			0.922			0.904
No	465 (96.67)	203 (97.13)		204 (97.61)	203 (97.13)	
Yes	9 (1.87)	3 (1.44)		3 (1.44)	3 (1.44)	
Unknown	7 (1.46)	3 (1.44)		2 (0.96)	3 (1.44)	
Lung metastases			*			0.615
No	462 (96.05)	208 (99.52)		206 (98.56)	208 (99.52)	
Yes	13 (2.70)	1 (0.48)		3 (1.44)	1 (0.48)	
Unknown	6 (1.25)	0 (0.00)		0 (0.00)	0 (0.00)	
Brain metastases			**			1
No	477 (99.17)	209 (100.00)		209 (100.00)	209 (100.00)	
Yes	0 (0.00)	0 (0.00)		0 (0.00)	0 (0.00)	
Unknown	4 (0.83)	0 (0.00)		0 (0.00)	0 (0.00)	

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

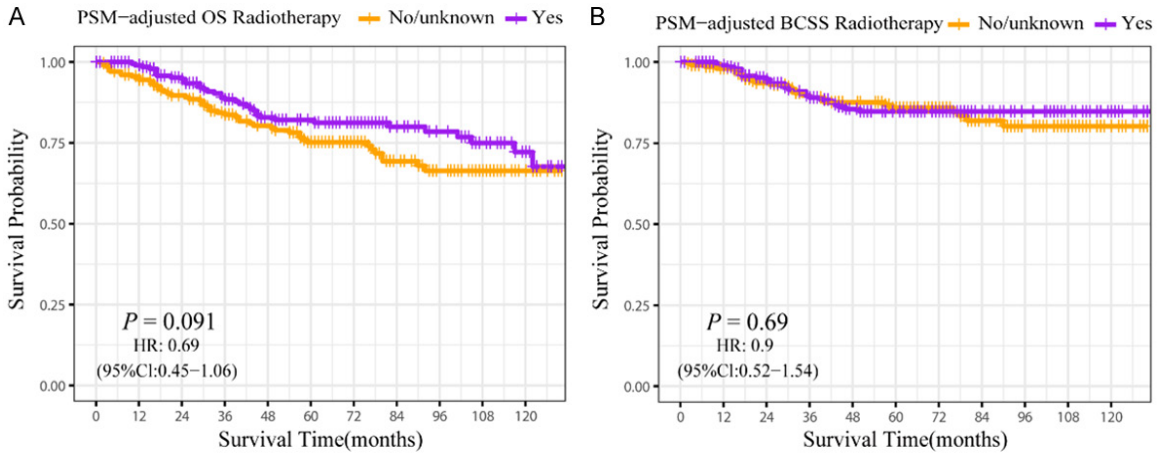


Figure 9. OS and BCSS of patients with PD-IDC who underwent radiotherapy after PSM adjustment. PSM-adjusted Kaplan-Meier (K-M) survival analysis: (A) OS of patients with PD-IDC who underwent radiotherapy versus those who did not; (B) BCSS of patients with PD-IDC underwent radiotherapy versus those who did not. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching.

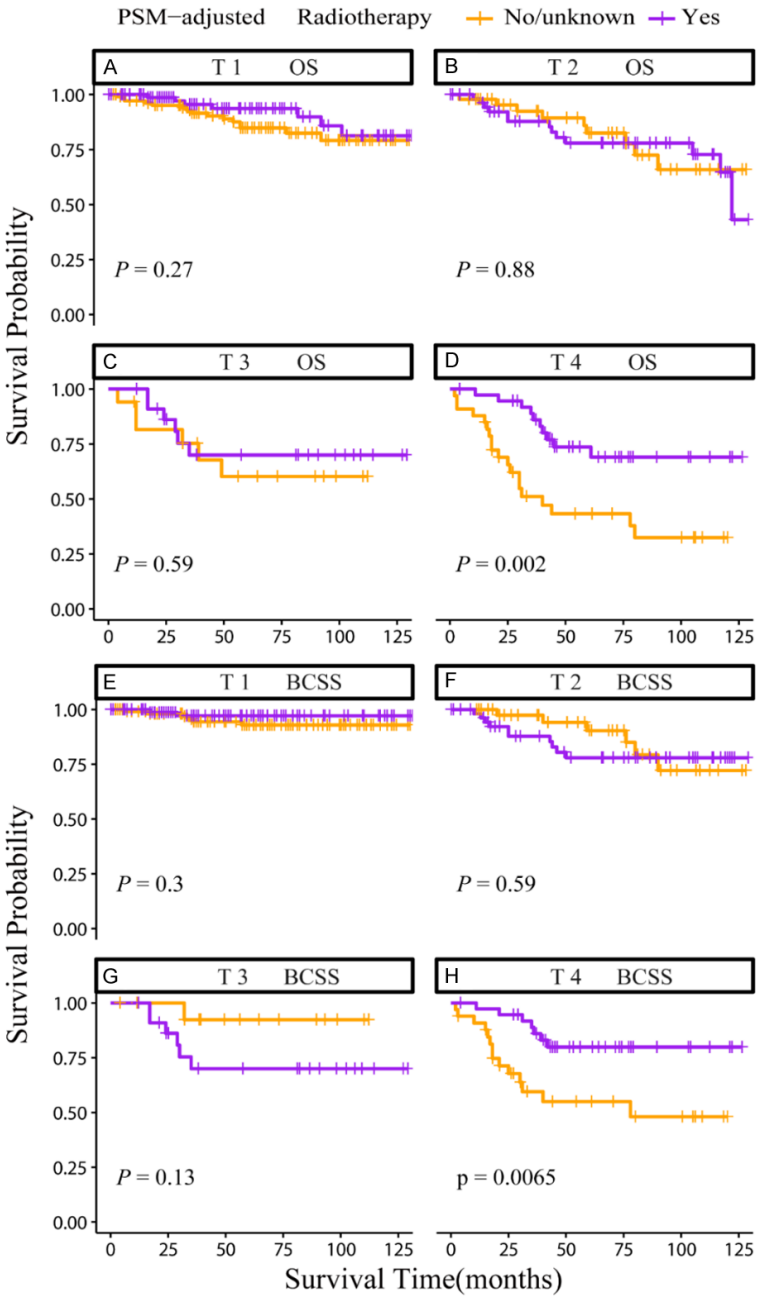


Figure 10. OS and BCSS of patients with PD-IDC who underwent radiotherapy after PSM adjustment (Stratified by stage T). A. OS of patients with PD-IDC with stage T1 tumors; B. OS of patients with PD-IDC with stage T2 tumors; C. OS of patients with PD-IDC with stage T3 tumors; D. OS of patients with PD-IDC with stage T4 tumors; E. BCSS of patients with PD-IDC with stage T1 tumors; F. BCSS of patients with PD-IDC with stage T2 tumors; G. BCSS of patients with PD-IDC with stage T3 tumors; H. BCSS of patients with PD-IDC with stage T4 tumors. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching.

tus was associated with a worse OS. However, a previous study suggested that unmarried status acted as a protective factor for patients with PD-IDC [10]. These contradictions may arise from the fact that the previous study did not further stratify the unmarried status or possibly from a misinterpretation of the reference factor. Typically, the HR-/HER-2- subtype of IDC is associated with the poorest prognosis. However, in our cohort of patients with PD-IDC, molecular type did not emerge as an independent prognostic factor, which contrasts with some previous studies [10, 19] but aligns with one particular study [8]. This variation could be attributed to the diverse patient populations included in these studies. Furthermore, we explore the impact of the time interval between diagnosis and therapy initiation as well as family income, which had not been previously investigated in patients with PD-IDC. Our results revealed that a high family income (> 74,999 *vs.* < 45,000 USD) could improve BCSS but did not affect the OS. Conversely, the time interval between diagnosis and therapy initiation did not prove to be a prognostic factor for PD-IDC. Interestingly, despite the generally poorer survival rates associated with PD-IDC, none of our patients had brain metastases at the time of diagnosis. Moreover, none of the previous studies had considered distant metastasis as a factor in univariable and multivariable Cox regression analyses for patients with PD-IDC. Our study was the first to incorporate these factors when conducting a Cox regression analysis for this patient population.

In our study, surgery, chemotherapy, and radiotherapy were not independent prognostic factors for patients with PD-IDC. It is worth noting

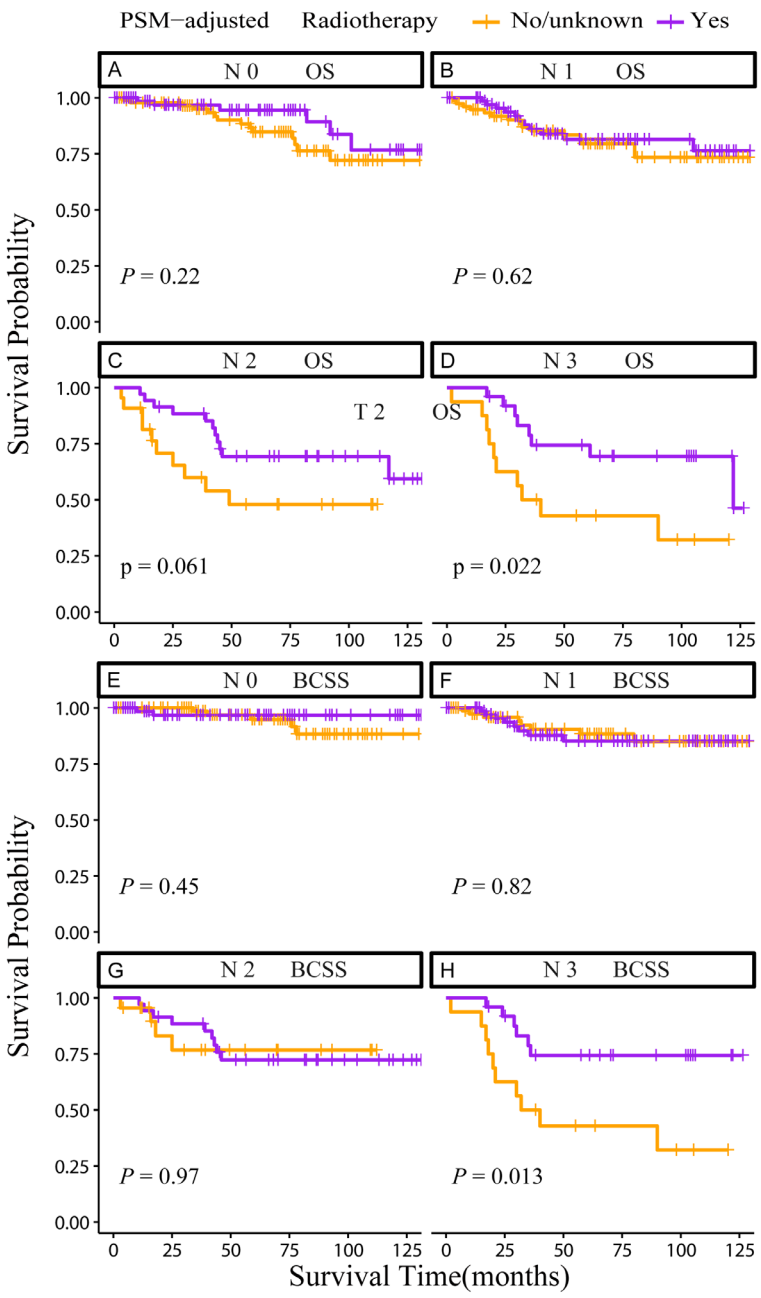


Figure 11. OS and BCSS of patients with PD-IDC who underwent radiotherapy after PSM adjustment (Stratified by stage N). A. OS of patients with PD-IDC with stage N0 lymph node involvement; B. OS of patients with PD-IDC with stage N1 lymph node involvement; C. OS of patients with PD-IDC with stage N2 lymph node involvement; D. OS of patients with PD-IDC with stage N3 lymph node involvement; E. BCSS of patients with PD-IDC with stage N0 lymph node involvement; F. BCSS of patients with PD-IDC with stage N1 lymph node involvement; G. BCSS of patients with PD-IDC with stage N2 lymph node involvement; H. BCSS of patients with PD-IDC with stage N3 lymph node involvement. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching.

that nearly all patients (96.38%) in our study underwent surgery, which might explain this result. Some studies have suggested that BCS followed by radiotherapy could be a viable treatment option [18, 19] for patients with PD-IDC when compared to mastectomy. Another recent meta-analysis reinforced the notion that, except for cases of isolated mammary PD, BCS alone is not recommended for treating PD-IDC and PD-DCIS [17], highlighting the importance of radiotherapy following BCS. Although chemotherapy and radiotherapy did not emerge as independent prognostic factors in our study, they remain pivotal treatments for breast cancer. Moreover, no prior study has analyzed the factors influencing the efficacy of chemotherapy and radiotherapy for patients with PD-IDC, prompting us to conduct further stratified analysis following PSM. Our findings indicated that chemotherapy failed to improve the prognosis of patients with PD-IDC across all molecular subtypes. However, within the grade III/IV subgroup, it significantly extended both OS and BCSS, emphasizing that it is the pathological grade, rather than the molecular subtype, that determines the efficacy of chemotherapy in improving the prognosis for patients with PD-IDC. Furthermore, our analysis revealed that only patients with PD-IDC with T4 and N3 stages experienced benefits from radiotherapy in terms of both OS and BCSS. However, this should not diminish the importance of radiotherapy in other subgroups, as its primary function is to reduce the local recurrence rates following mastectomy, BCS, and axillary lymph node dissection [31, 32]. In summary, our findings underscore the crucial role of radiotherapy

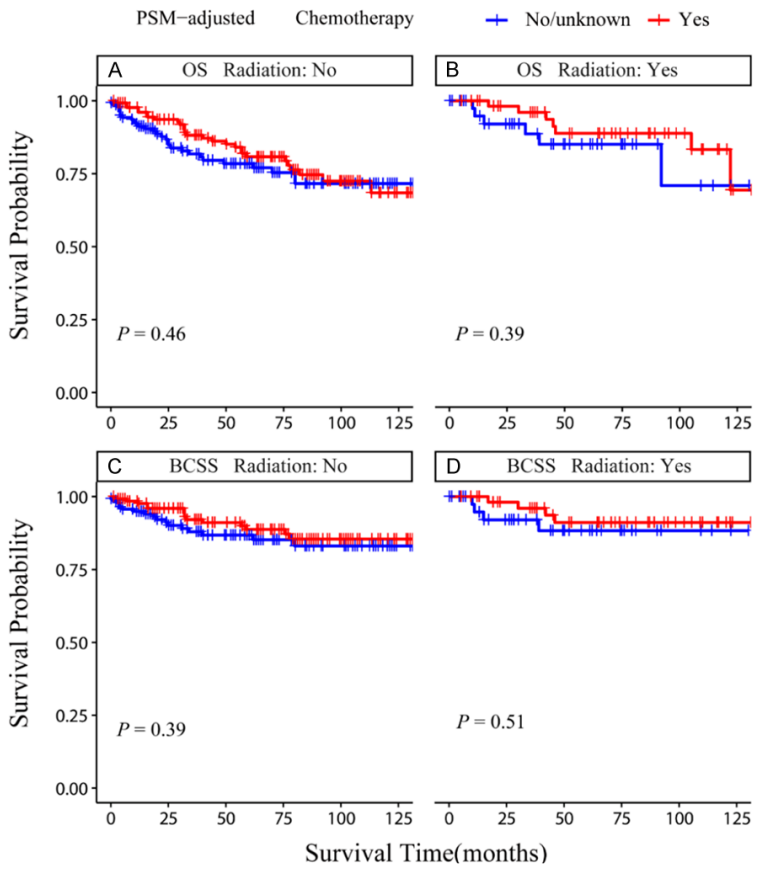


Figure 12. OS and BCSS of PD-IDC patients underwent chemotherapy or not after PSM adjustment (Stratified by radiotherapy). A. OS of PD-IDC patients without radiotherapy; B. OS of PD-IDC patients with radiotherapy; C. BCSS of PD-IDC patients without radiotherapy; D. BCSS of PD-IDC patients with radiotherapy.

for patients with PD-IDC with T4 and N3 stages, as it significantly contributes to their survival.

There are still some limitations in our study, despite its promising discoveries. First, nearly all the patients (96.38%) underwent surgery, making it hard to assess whether patients with PD-IDC could be exempt from surgery in some cases. Second, the SEER data may be a good representation of the general situation, but because of ethnic differences, it may not always be the case for Asians and the Chinese in particular. Third, due to the limited number of cases, the number of matches in the PSM was not always 100%, so selection bias may have occurred. In the end, because there is no information on endocrine therapy in this version of the SEER database, we cannot investigate the role of endocrine therapy in patients with PD-IDC.

Conclusions

We comprehensively analyzed the clinical characteristics and prognosis of patients with PD-IDC and developed a user-friendly web-based nomogram to predict their survival. This predictive model not only demonstrates high accuracy but also offers a more accessible tool for healthcare providers and patients compared to traditional nomograms. Furthermore, our findings from the in-depth stratified analysis underscored a crucial point: it is the pathological grade, rather than the molecular subtype, that determines the efficacy of chemotherapy in improving the prognosis for patients with PD-IDC. Additionally, our study highlighted the significance of radiotherapy in prolonging the survival of patients with PD-IDC specifically in cases involving T4 and N3 stages.

Acknowledgements

This study was supported by the following: Free exploring fund of Xi'an Jiaotong University (xzy012022-096, to X.X. Zhao; xzy012022097, to J.K. Qu). Medical "basic -clinical" integration and innovation project (YXJLRH2022088, to J.K. Qu). Key laboratory construction project (2022-ZXY-SYS-002, to S.Q. Zhang).

Disclosure of conflict of interest

None.

Address correspondence to: Shuqun Zhang, Department of Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 West Fifth Street, Xi'an, Shaanxi, P. R. China. E-mail: zhangshuqun0001@mail.xjtu.edu.cn; Xixi Zhao, Department of Radiation Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 West Fifth Street, Xi'an, Shaanxi, P. R. China. E-mail: xixizhao@xjtu.edu.cn

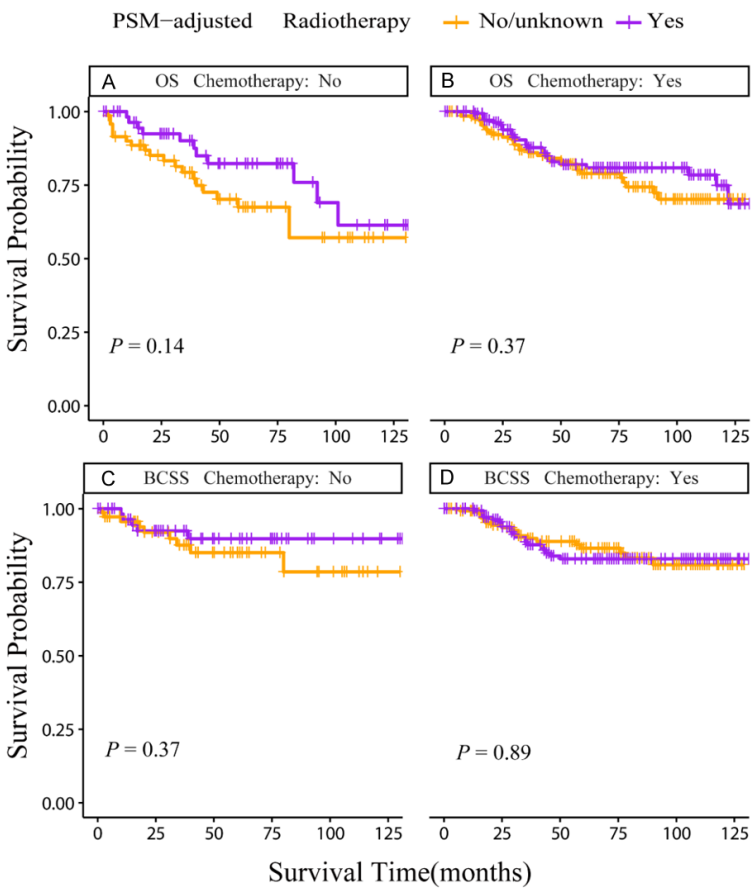


Figure 13. OS and BCSS of PD-IDC patients underwent radiotherapy or not after PSM adjustment (Stratified by chemotherapy). A. OS of PD-IDC patients without chemotherapy; B. OS of PD-IDC patients with chemotherapy; C. BCSS of PD-IDC patients without chemotherapy; D. BCSS of PD-IDC patients with chemotherapy.

References

[1] Chen CY, Sun LM and Anderson BO. Paget disease of the breast: changing patterns of incidence, clinical presentation, and treatment in the U.S. Cancer 2006; 107: 1448-1458.

[2] Caliskan M, Gatti G, Sosnovskikh I, Rotmensz N, Botteri E, Musmeci S, Rosali dos Santos G, Viale G and Luini A. Paget's disease of the breast: the experience of the European Institute of Oncology and review of the literature. Breast Cancer Res Treat 2008; 112: 513-521.

[3] Markarian S and Holmes DR. Mammary Paget's disease: an update. Cancers (Basel) 2022; 14: 2422.

[4] Li H, Xiao Y, Li Q, Yao J, Yuan X, Zhang Y, Yin X, Saito Y, Fan H, Li P, Kuo WL, Halpin A, Gibbons DL, Yagita H, Zhao Z, Pang D, Ren G, Yee C, Lee JJ and Yu D. The allergy mediator histamine confers resistance to immunotherapy in cancer patients via activation of the macrophage histamine receptor H1. Cancer Cell 2022; 40: 36-52, e9.

[5] Choridah L, Sari WK, Dwiangingsih EK, Widodo I, Suwardjo and Anwar SL. Advanced lesions of synchronous bilateral mammary Paget's disease: a case report. J Med Case Rep 2020; 14: 119.

[6] Sandoval-Leon AC, Drews-Elger K, Gomez-Fernandez CR, Yepes MM and Lippman ME. Paget's disease of the nipple. Breast Cancer Res Treat 2013; 141: 1-12.

[7] Wong SM, Freedman RA, Sagara Y, Stamell EF, Desantis SD, Barry WT and Golshan M. The effect of Paget disease on axillary lymph node metastases and survival in invasive ductal carcinoma. Cancer 2015; 121: 4333-4340.

[8] Chen S, Chen H, Yi Y, Jiang X, Lei H, Luo X, Chen Y, Liu S, Yuan D, Jia X and Li J. Comparative study of breast cancer with or without concomitant Paget disease: an analysis of the SEER database. Cancer Med 2019; 8: 4043-4054.

[9] Zhou H, Lu K, Zheng L, Guo L, Gao Y, Miao X, Chen Z and Wang X. Prognostic significance of mammary Paget's disease in Chinese women: a 10-year, population-based, matched cohort study. Onco Targets Ther 2018; 11: 8319-8326.

[10] Zhao Y, Sun HF, Chen MT, Gao SP, Li LD, Jiang HL and Jin W. Clinicopathological characteristics and survival outcomes in Paget disease: a SEER population-based study. Cancer Med 2018; 7: 2307-2318.

[11] Marshall JK, Griffith KA, Haffty BG, Solin LJ, Vicini FA, McCormick B, Wazer DE, Recht A and Pierce LJ. Conservative management of Paget disease of the breast with radiotherapy: 10- and 15-year results. Cancer 2003; 97: 2142-2149.

[12] Karakas C. Paget's disease of the breast. J Carcinog 2011; 10: 31.

[13] Chen X, Zhang C, Guo D, Wang Y, Hu J, Hu J, Wang S and Liu X. Distant metastasis and prognostic factors in patients with invasive ductal carcinoma of the breast. Eur J Clin Invest 2022; 52: e13704.

[14] Feng Y, Zhang Y, Xiang Y, Guo K, Jin H, Ruan S and Guan Z. Nomograms for predicting specific distant metastatic sites and overall survival of breast invasive ductal carcinoma patients after surgery: a large population-based study. Front Surg 2022; 9: 779220.
[15] Dai D, Jin H and Wang X. Nomogram for predicting survival in triple-negative breast cancer patients with histology of infiltrating duct carcinoma: a population-based study. Am J Cancer Res 2018; 8: 1576-1585.
[16] Pelorca RJF, de Oliveira-Junior I and da Costa Vieira RA. Oncoplastic surgery for Paget's disease of the breast. Front Oncol 2023; 13: 1151932.
[17] Lin CW, Chiang MH and Tam KW. Treatment of mammary Paget disease: a systematic review and meta-analysis of real-world data. Int J Surg 2022; 107: 106964.
[18] Yao Y, Sun L, Meng Y, Zhuang Y, Zhao L, Yu Q and Si C. Breast-conserving surgery in patients with mammary Paget's disease. J Surg Res 2019; 241: 178-187.
[19] Wu Q, Ding X, Li J, Sun S, Zhu S, Wu J, Liu Q, Yao F and Sun S. Surgical treatment in Paget's disease with invasive ductal carcinoma: an observational study based on SEER. Sci Rep 2017; 7: 45510.
[20] Helme S, Harvey K and Agrawal A. Breast-conserving surgery in patients with Paget's disease. Br J Surg 2015; 102: 1167-1174.
[21] Wong SM, Freedman RA, Stamell E, Sagara Y, Brock JE, Desantis SD and Golshan M. Modern trends in the surgical management of Paget's disease. Ann Surg Oncol 2015; 22: 3308-3316.
[22] Kane LT, Fang T, Galetta MS, Goyal DKC, Nicholson KJ, Kepler CK, Vaccaro AR and Schroeder GD. Propensity score matching: a statistical method. Clin Spine Surg 2020; 33: 120-122.
[23] Barakat A, Mittal A, Ricketts D and Rogers BA. Understanding survival analysis: actuarial life tables and the Kaplan-Meier plot. Br J Hosp Med (Lond) 2019; 80: 642-646.
[24] Austin PC, Putter H, Giardiello D and van Klaveren D. Graphical calibration curves and the integrated calibration index (ICI) for competing risk models. Diagn Progn Res 2022; 6: 2.
[25] Vickers AJ, Cronin AM, Elkin EB and Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 2008; 8: 53.
[26] Jarząb M, Stobiecka E, Badora-Rybicka A, Chmielik E, Kowalska M, Bal W, Polakiewicz-Gilowska A, Bobek-Billewicz B, Lange D and Tarnawski R. Association of breast cancer grade with response to neoadjuvant chemotherapy assessed postoperatively. Pol J Pathol 2019; 70: 91-99.
[27] Guo Z, Li W, Wu K, Fu Y, Yan R and Zhou X. A nomogram based on the log odds of positive lymph nodes for predicting the prognosis of T1 stage rectal cancer. Am J Cancer Res 2023; 13: 1498-1508.
[28] Yang HS, Li B, Liu SH and Ao M. Nomogram model for predicting postoperative survival of patients with stage IB-IIA cervical cancer. Am J Cancer Res 2021; 11: 5559-5570.
[29] Tao Q, Zeng Q, Liu W, Liu J, Jiang L, Tu X, Li K, Zhao P, Tang X, Liu Z, Wang L, Xu Q and Zheng Y. A novel prognostic nomogram for hepatocellular carcinoma after thermal ablation. Am J Cancer Res 2021; 11: 5126-5140.
[30] Ma Z, Huang B, Huang S, Liu C, Cao J, Zheng Z, Li Z, Zhou Z, Zhuang H, Zou Y, Yang L, Guo J, Zhang C and Hou B. Prognostic stratification based on a novel nomogram for left-sided pancreatic adenocarcinoma after surgical resection: a multi-center study. Am J Cancer Res 2021; 11: 2754-2768.
[31] Piras A, Boldrini L, Venuti V, Sanfratello A, La Vecchia M, Gennari R, Sortino G, Angileri T and Daidone A. Mammary Paget's disease and radiotherapy: a systematic literature review. Eur Rev Med Pharmacol Sci 2021; 25: 1821-1827.
[32] Zhang JQ, Lu CY, Qin L, Chen HM and Wu SY. Outcome of post-mastectomy radiotherapy after primary systemic treatment in patients with different clinical tumor and nodal stages of breast cancer: a cohort study. Am J Cancer Res 2020; 10: 2185-2198.

Supplementary Table 1. Univariate and multivariate Cox analysis of PD-IDC characteristics extracted from train data

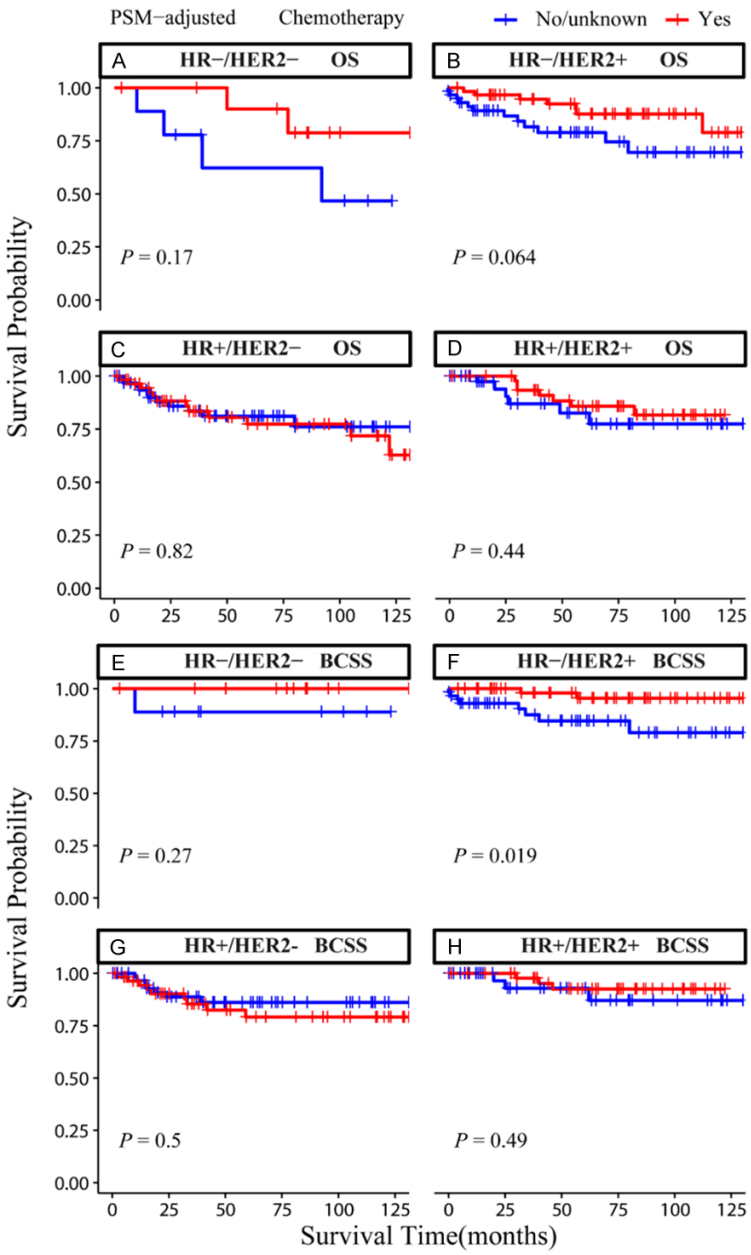
	OS					
	Univariate cox analysis			Multivariate cox analysis		
	HR	95% CI	*P* Value	HR	95% CI	*P* Value
Age						
24~90	1.07	1.05-1.08	***	1.05	1.02-1.07	***
Marriage status						
Married	reference			reference		
Divorced/Separated	0.87	0.38-1.98	0.74	1.11	0.45-2.70	0.83
Single	2.89	1.71-4.88	***	2.14	1.09-4.22	*
Widowed	5.75	3.54-9.32	***	2.24	1.14-4.38	*
Race						
White	reference			reference		
Black	1.12	0.62-2.00	0.71	/	/	/
Other	0.67	0.32-1.37	0.27	/	/	/
Months from diagnosis to therapy						
0 month	reference			reference		
≥1 month	0.88	0.58-1.35	0.57	/	/	/
Median household income(inflation ajusted)						
<45,000$	reference			reference		
45,000-54,999$	0.35	0.15-0.83	*	0.57	0.17-1.90	0.36
55,000-64,999$	0.48	0.24-0.94	*	0.64	0.23-1.80	0.40
65,000-74,999$	0.38	0.20-0.75	**	0.69	0.25-1.93	0.48
>74,999$	0.27	0.14-0.54	***	0.56	0.20-1.60	0.28
Grade						
Grade I/Grade II	reference			reference		
Grade III/Grade IV	1.73	1.12-2.68	*	1.96	1.06-3.64	*
Subtypes						
HR-/HER2-	reference			reference		
HR-/HER2+	0.28	0.14-0.58	***	0.51	0.19-1.40	0.19
HR+/HER2-	0.50	0.25-0.99	*	0.61	0.21-1.76	0.36
HR+/HER2+	0.65	0.34-1.26	0.21	0.98	0.39-2.47	0.96
T Stage						
T1	reference			reference		
T2	1.81	1.10-2.96	*	1.25	0.64-2.45	0.52
T3	2.63	1.27-5.5	**	2.79	0.94-8.29	0.066
T4	4.12	2.57-6.62	***	3.21	1.57-6.59	**
N Stage						
N0	reference			reference		
N1	1.56	0.99-2.46	0.05	1.77	0.96-3.28	0.068
N2	2.66	1.49-4.73	***	2.14	0.93-4.91	0.072
N3	2.77	1.51-5.08	***	1.94	0.68-5.54	0.21
Treatment combination (Radiation or Chemotherapy)						
No/Unknown	reference			reference		
Radiation only	0.42	0.19-0.91	*	0.84	0.32-2.22	0.72
Chemotherapy only	0.36	0.23-0.58	***	0.48	0.22-1.05	0.07
Radiation and Chemotherapy	0.34	0.19-0.59	***	0.49	0.21-1.16	0.10

Surgery						
No	reference			reference		
Yes	0.35	0.15-0.80	*	0.76	0.14-4.08	0.74
Bone metastasis						
No	reference					
Yes	7.20	3.56-14.39	***	2.71	0.65-11.24	0.17
Liver metastasis						
No	reference			reference		
Yes	6.83	2.97-15.70	***	3.22	0.67-15.59	0.15
Lung metastasis						
No	reference			reference		
Yes	8.91	4.10-19.36	***	2.52	0.69-9.15	0.16

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

Supplementary Table 2. Relevant characteristics of PD-IDC patients included from our and sister hospitals for Nomogram external validation

Age	Marital	Grade	T	N	Bone
41	Married	Grade III/Grade IV	T2	N0	No
54	Married	Grade I/II	T1	N0	No
55	Married	Grade I/II	T1	N0	No
45	Married	Grade I/II	T2	N1	No
75	Married	Grade I/II	T1	N0	No
52	Married	Grade I/II	T2	N1	No
51	Married	Grade I/II	T1	N1	No
66	Married	Grade I/II	T2	N0	No
46	Married	Grade I/II	T1	N1	No
85	Married	Grade III/Grade IV	T3	N3	Yes
82	Widowed	Grade III/Grade IV	T4	N3	No
83	Widowed	Grade III/Grade IV	T3	N3	No



Supplementary Figure 1. OS and BCSS of PD-IDC patients underwent chemotherapy or not after PSM adjustment (Stratified by molecular subtype). A. OS of PD-IDC patients with HR-Her2-; B. OS of PD-IDC patients with HR-Her2+; C. OS of PD-IDC patients with HR+Her2-; D. OS of PD-IDC patients with HR+Her2+; E. BCSS of PD-IDC patients with HR-Her2-; F. BCSS of PD-IDC patients with HR-Her2+; G. BCSS of PD-IDC patients with HR+Her2-; H. BCSS of PD-IDC patients with HR+Her2+. OS, overall survival; BCSS, breast cancer-specific survival; PD-IDC, Paget's disease-invasive ductal carcinoma; PSM, propensity score matching; HR, hormone receptor; Her-2, human epidermal growth factor receptor 2.