Original Article

A nomogram for predicting brain metastasis in patients with de novo stage IV breast cancer

Ming-Shuai Sun^, Yin-Hua Liu, Jing-Ming Ye, Qian Liu, Yuan-Jia Cheng, Ling Xin, Ling Xu^

Breast Disease Center, Peking University First Hospital, Beijing, China

Contributions: (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: MS Sun, L Xu; (V) Data analysis and interpretation: MS Sun; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Ling Xu. Breast Disease Center, Peking University First Hospital, Beijing 100034, China. Email: ling.xu@pkufh.com.

Background: Brain metastasis (BM) is a very serious event in patients with breast cancer. The aim of this study was to establish a nomogram to predict the risk of BM in patients with *de novo* stage IV breast cancer.

Methods: We gathered female patients diagnosed with de novo stage IV breast cancer between 2010 and 2015 from the Surveillance, Epidemiology, and End Results (SEER) database. After randomly allocating the patients to the training set and verification set, we used univariate and multivariate logistic regression to analyze the relationship between BM and clinicopathological features. Finally, we developed a nomogram which was validated by the analysis of calibration curve and receiver operating characteristic curve.

Results: Of 7,154 patients with *de novo* stage IV breast cancer, 422 developed BM. Age, tumor size, subtype, and the degree of lung involvement were significantly correlated with BM. The nomogram had discriminatory ability with an area under curve (AUC) of 0.640 [95% confidence interval (CI): 0.607 to 0.673] in the training set, and 0.644 (95% CI: 0.595 to 0.693) in the validation set.

Conclusions: Our study developed a nomogram to predict BM for de novo stage IV breast cancer, thus helping clinicians to identify patients at high-risk of BM and implement early preventive interventions to improve their prognoses.

Keywords: Breast cancer (BC); brain metastasis (BM); nomogram; Surveillance, Epidemiology, and End Results (SEER)

Submitted Mar 10, 2021. Accepted for publication May 13, 2021.
doi: 10.21037/atm-21-1808
View this article at: http://dx.doi.org/10.21037/atm-21-1808

Introduction

Breast cancer tumors have the highest incidence of malignancy in women. The overall prognosis of patients with breast cancer in the United States is good, with a 5-year survival rate of 90.2% (1). However, the prognosis of metastatic breast cancer (MBC) is very poor, and brain metastasis (BM) is one of the distant metastatic sites with the worst prognosis. Although the incidence of BM in early breast cancer is less than 3% (2,3), the incidence in patients with MBC is as high as 10–30% (3-5), and the incidence in autopsy reports of patients with MBC exceeds 30% (6). Patients with BM have rapid disease progression, poor quality of life, high mortality, and few effective treatment options, resulting in a median survival time of only 4 weeks for untreated breast cancer patients with BM (6,7).

In recent years, remarkable progress has been made in the treatment of BM, novel targeted therapeutic drugs

^ ORCID: Ming-Shuai Sun, 0000-0003-2645-9568; Ling Xu, 0000-0002-9236-3014.

and locoregional treatment techniques have brought survival benefits to patients with BM, especially in human epidermal growth factor receptor 2 (HER2) positive breast cancer. Small molecule tyrosine kinase inhibitors (TKI) such as tucatinib (8,9) and neratinib (10-12), as well as novel antibody-drug conjugates such as trastuzumab emtansine (T-DM1) (13,14) and trastuzumab deruxtecan (15,16), have shown substantial intracranial efficacy against HER2-positive breast cancer. The application of new locoregional treatments, such as CyberKnife radiosurgery and stereotactic radiotherapy, has significantly improved the prognosis of patients with BM (17). The median survival time of breast cancer patients with BM treated with whole brain radiotherapy and stereotactic radiosurgery can be extended to 4–6 months, and if a single metastatic focus can be surgically removed, the median survival time can be extended to 16 months (6,7,18). Nevertheless, the overall prognosis of BM remains poor, mainly due to late detection. Therefore, early identification of patients with high-risk of BM is the key to improving their outcomes. Clinicians can perform routine laboratory or imaging screening for high-risk patients, which is helpful for early detection, early treatment and even preventive intervention, thus improving the prognosis of these patients.

However, no proven screening method has currently been established for BM. For screening BM in patients with early breast cancer, only routine central nervous system (CNS) symptoms follow up is recommended, no routine laboratory or imaging screening is recommended for patients in the absence of CNS symptoms due to a lack of research confirming associated survival advantages (19). This approach is even applicable to patients with advanced breast cancer, including metastatic HER2-positive and triple-negative breast cancer (TNBC) (20,21). Owing to the low incidence of BM in breast cancer, especially early breast cancer, regular imaging screening for all patients is not cost-effective, and the cumulative radiation is detrimental for the body. In view of this, MRI screening is needed only when patients have CNS symptoms, but a high proportion of patients do not have CNS symptoms in the early stage of BM (6), so the screening model recommended by the guidelines cannot distinguish BM in time. Therefore, researchers have tried to identify pathological and biological parameters associated with BM. However, no highly abundant and recurrent mutations have been proved to be associated with BM in breast cancer, even the ERBB2-activated breast cancer genetically engineered mouse models did not interrogate the presence of brain metastases (22). Previous studies have only found that MBC patients with HER2-positive breast cancer (23), TNBC (24,25), *BRCA* genes mutation (26,27), diagnosis at young age (28), lung metastasis (29), high histological grade (23), and high proliferative activity (30) are more likely to develop BM.

In our study, we incorporated a large number of MBC cases from the Surveillance, Epidemiology, and End Results (SEER) database with 2 goals in mind. The primary goal was to identify risk factors associated with BM, and the secondary goal was to establish a predictive model for BM and evaluate its performance by examining its correlation with survival and designing a preventive intervention trial. The predictive model established in this study was expected to provide a basis for BM screening and preventive treatment in patients with MBC.

We present the following article in accordance with the TRIPOD reporting checklist (available at http://dx.doi.org/10.21037/atm-21-1808).

Methods

Data source and inclusion criteria

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). We searched the SEER database released in November 2018 which covers approximately 30% of the U.S. population from 18 registries (1973–2015). Informed consent is not required to extract SEER data because it does not provide identification information. The inclusion criteria were as follows: (I) primary malignant breast cancer diagnosed with distant metastasis from 2010 to 2015, patients before 2010 were excluded because their HER2 statuses were not recorded in the SEER database; (II) female, age at diagnosis ≥18; (III) histologically diagnosed as invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), or IDC or ILC mixed with other types of carcinoma; (IV) breast cancer was the only primary malignant tumor; (V) patients who were not diagnosed by autopsy or death certificate; (VI) patients with complete clinicopathological data, those with missing data were excluded. The brain involvements of all patients were assessed at the time of initial diagnosis, so the patients had not received any form of treatment for breast cancer before. The evaluation of all the variables in the study was for the purpose of routine diagnosis and treatment, and the data were collected retrospectively when constructing the nomogram, therefore there was no obvious assessment bias. All the variable definitions in the study can

be found in the SEER program coding and staging manual (https://seer.cancer.gov/tools/codingmanuals/). In our study, after excluding patients whose clinicopathological information was not available, we identified 7,154 patients, of whom 422 (5.9%) had clinical evidence of BM, which was regarded as an outcome event. Patients were randomly allocated to a training set and validation set according to the proportion of 7:3, the training set was used to establish the prediction model of BM, and the validation set was used for validation of the model.

Statistical analysis

Pearson chi-square test was used to evaluate the baseline characteristics and test the relationship between BM and clinicopathologic characteristics. Variables that were found to be statistically significant in the Pearson chi-square test were incorporated into multivariate logistic regression analysis. Significant variables in multivariate logistic regression were considered independent predictors of BM and were included in the final multivariate logistic regression model (nomogram).

The nomogram performance was evaluated from the aspects of discrimination, calibration, and clinical utility. The discrimination was quantified with the area under curve (AUC). The calibration was evaluated by the calibration curve using the bootstrapping method (1,000 repeats), which represented the relationship between observation frequencies and prediction probabilities. In order to evaluate the clinical utility of the nomogram, we defined the predictive risk of BM as a new variable (above/ below mean) and performed Cox regression analysis together with other confirmed independent prognostic factors of breast cancer (including liver involvement, surgery, chemotherapy, and so on). Multivariate cox regression analysis was aimed to reveal this comprehensive BM prediction risk as one of the most important variables for predicting survival outcomes. The clinical utility of this nomogram was also demonstrated by conducting a virtual trial in the validation set to prevent BM. We used the nomogram to calculate the individual risk of BM in the validation set and set different thresholds as intervention conditions. We evaluated the health economic value of our model in the prevention of breast cancer BM by estimating the number of patients who needed preventive intervention and the number of BM successfully prevented.

We compared the overall survival (OS) between two groups by Kaplan-Meier curves and log-rank test. All of the statistical methods in our study were performed using the statistical software SPSS version 23 (IBM Corp., Chicago, IL, USA) and R software version 3.6.3 (https://www.R-project.org/). All statistical analyses were 2-sided, and statistical significance was defined at $P<0.05$. All confidence intervals (CIs) were reported at the 95% confidence level.

Results

Demographic characteristics

Based on the inclusion criteria, we recruited a cohort of 7,154 patients with breast cancer with distant metastasis from the SEER database, among whom 422 patients (5.9%) had clinical evidence of BM. *Table 1* shows the clinicopathologic characteristics of the entire cohort, as well as the training and validation sets which were randomly derived from the entire cohort at a 7:3 ratio. We used Pearson chi-square test for the training set and demonstrated that the status of brain involvement was related to variables including age at diagnosis ($P=0.038$), grade ($P=0.043$), histological type ($P=0.027$), tumor size ($P=0.034$), estrogen receptor (ER) status ($P<0.001$), progesterone receptor (PR) status ($P<0.001$), subtype ($P<0.001$), degree of liver involvement ($P=0.004$), and degree of lung involvement ($P<0.001$).

Multivariate logistic regression results

Variables that were found to be statistically significant in the Pearson chi-square test were incorporated into multivariate logistic regression analysis in *Table 2*. In multivariate logistic regression analysis, BM was independently associated with age at diagnosis, tumor size, subtype, and the degree of lung involvement.

Nomogram construction

We used the results of the multivariate logistic regression in *Table 2* to construct a nomogram to predict the risk of BM (*Figure 1*). By summing the scores of each variable, we can predict the probability of BM in a specific patient. Interestingly, the relationship between age and BM was neither a straightforward positive nor a negative correlation. The results showed that patients aged 45–64 had the highest risk of BM, followed by patients over 64 years old, and the risk was lowest for patients under 45 years old. Similarly, tumors smaller than 2 cm had the highest risk of BM, followed by tumors over 5 cm, and those sized >2 cm and ≤5 cm had the lowest risk.

Table 1 Participant characteristics for the whole cohort, training set, and validation set

Variables	Whole cohort				Training set				Validation set			
	All, N=7,154	Non-brain metastasis, N=6,732	Brain metastasis, N=422	P value	All, N=5,007	Non-brain metastasis, N=4,712	Brain metastasis, N=295	P value	All, N=2,147	Non-brain metastasis, N=2,020	Brain metastasis, N=127	P value
Age				0.001				0.038				0.005
<45	1,112 (15.5%)	1,059 (15.7%)	53 (12.6%)		797 (15.9%)	760 (16.1%)	37 (12.5%)		315 (14.7%)	299 (14.8%)	16 (12.6%)	
45–64	3,529 (49.3%)	3,283 (48.8%)	246 (58.3%)		2,451 (49.0%)	2,286 (48.5%)	165 (55.9%)		1,078 (50.2%)	997 (49.4%)	81 (63.8%)	
>64	2,513 (35.1%)	2,390 (35.5%)	123 (29.1%)		1,759 (35.1%)	1,666 (35.4%)	93 (31.5%)		754 (35.1%)	724 (35.8%)	30 (23.6%)	
Race				0.664				0.922				0.244
Black	1,232 (17.2%)	1,163 (17.3%)	69 (16.4%)		868 (17.3%)	815 (17.3%)	53 (18.0%)		364 (17.0%)	348 (17.2%)	16 (12.6%)	
White	5,305 (74.2%)	4,993 (74.2%)	312 (73.9%)		3,699 (73.9%)	3,484 (73.9%)	215 (72.9%)		1,606 (74.8%)	1,509 (74.7%)	97 (76.4%)	
Others	617 (8.62%)	576 (8.56%)	41 (9.72%)		440 (8.79%)	413 (8.76%)	27 (9.15%)		177 (8.24%)	163 (8.07%)	14 (11.0%)	
Marital status				0.256				0.123				0.846
Unmarried	3,750 (52.4%)	3,517 (52.2%)	233 (55.2%)		2,591 (51.7%)	2,425 (51.5%)	166 (56.3%)		1,159 (54.0%)	1,092 (54.1%)	67 (52.8%)	
Married	3,404 (47.6%)	3,215 (47.8%)	189 (44.8%)		2,416 (48.3%)	2,287 (48.5%)	129 (43.7%)		988 (46.0%)	928 (45.9%)	60 (47.2%)	
Laterality				0.433				0.536				0.693
Left	3,667 (51.3%)	3,459 (51.4%)	208 (49.3%)		2,557 (51.1%)	2,412 (51.2%)	145 (49.2%)		1,110 (51.7%)	1,047 (51.8%)	63 (49.6%)	
Right	3,487 (48.7%)	3,273 (48.6%)	214 (50.7%)		2,450 (48.9%)	2,300 (48.8%)	150 (50.8%)		1,037 (48.3%)	973 (48.2%)	64 (50.4%)	
Primary site				0.009				0.179				0.013
Central	497 (6.95%)	481 (7.14%)	16 (3.79%)		344 (6.87%)	331 (7.02%)	13 (4.41%)		153 (7.13%)	150 (7.43%)	3 (2.36%)	
Inner	874 (12.2%)	832 (12.4%)	42 (9.95%)		608 (12.1%)	580 (12.3%)	28 (9.49%)		266 (12.4%)	252 (12.5%)	14 (11.0%)	
Outer	2,485 (34.7%)	2,338 (34.7%)	147 (34.8%)		1,753 (35.0%)	1,642 (34.8%)	111 (37.6%)		732 (34.1%)	696 (34.5%)	36 (28.3%)	
Overlap	1,609 (22.5%)	1,514 (22.5%)	95 (22.5%)		1,136 (22.7%)	1,070 (22.7%)	66 (22.4%)		473 (22.0%)	444 (22.0%)	29 (22.8%)	
Unknown	1,689 (23.6%)	1,567 (23.3%)	122 (28.9%)		1,166 (23.3%)	1,089 (23.1%)	77 (26.1%)		523 (24.4%)	478 (23.7%)	45 (35.4%)	
Grade				0.019				0.043				0.155
I	508 (7.10%)	489 (7.26%)	19 (4.50%)		356 (7.11%)	342 (7.26%)	14 (4.75%)		152 (7.08%)	147 (7.28%)	5 (3.94%)	
II	2,927 (40.9%)	2,767 (41.1%)	160 (37.9%)		2,081 (41.6%)	1,973 (41.9%)	108 (36.6%)		846 (39.4%)	794 (39.3%)	52 (40.9%)	
III	3,682 (51.5%)	3,443 (51.1%)	239 (56.6%)		2,544 (50.8%)	2,373 (50.4%)	171 (58.0%)		1,138 (53.0%)	1,070 (53.0%)	68 (53.5%)	
IV	37 (0.52%)	33 (0.49%)	4 (0.95%)		26 (0.52%)	24 (0.51%)	2 (0.68%)		11 (0.51%)	9 (0.45%)	2 (1.57%)	

Table 1 *(continued)*

Table 1 *(continued)*

Variables	Whole cohort				Training set				Validation set			
	All, N=7,154	Non-brain metastasis, N=6,732	Brain metastasis, N=422	P value	All, N=5,007	Non-brain metastasis, N=4,712	Brain metastasis, N=295	P value	All, N=2,147	Non-brain metastasis, N=2,020	Brain metastasis, N=127	P value
Histological type				0.013				0.027				0.341
IDC	6,038 (84.4%)	5,661 (84.1%)	377 (89.3%)		4,221 (84.3%)	3,956 (84.0%)	265 (89.8%)		1,817 (84.6%)	1,705 (84.4%)	112 (88.2%)	
ILC	628 (8.78%)	605 (8.99%)	23 (5.45%)		453 (9.05%)	436 (9.25%)	17 (5.76%)		175 (8.15%)	169 (8.37%)	6 (4.72%)	
Others	488 (6.82%)	466 (6.92%)	22 (5.21%)		333 (6.65%)	320 (6.79%)	13 (4.41%)		155 (7.22%)	146 (7.23%)	9 (7.09%)	
Tumor size (cm)				0.010				0.034				0.279
≤2	1,074 (15.0%)	1,001 (14.9%)	73 (17.3%)		743 (14.8%)	693 (14.7%)	50 (16.9%)		331 (15.4%)	308 (15.2%)	23 (18.1%)	
>2 and ≤5	3,412 (47.7%)	3,241 (48.1%)	171 (40.5%)		2,370 (47.3%)	2,252 (47.8%)	118 (40.0%)		1,042 (48.5%)	989 (49.0%)	53 (41.7%)	
>5	2,668 (37.3%)	2,490 (37.0%)	178 (42.2%)		1,894 (37.8%)	1,767 (37.5%)	127 (43.1%)		774 (36.1%)	723 (35.8%)	51 (40.2%)	
N stage				0.231				0.101				0.553
N0	1,512 (21.1%)	1,407 (20.9%)	105 (24.9%)		1,039 (20.8%)	963 (20.4%)	76 (25.8%)		473 (22.0%)	444 (22.0%)	29 (22.8%)	
N1	3,390 (47.4%)	3,196 (47.5%)	194 (46.0%)		2,401 (48.0%)	2,271 (48.2%)	130 (44.1%)		989 (46.1%)	925 (45.8%)	64 (50.4%)	
N2	976 (13.6%)	926 (13.8%)	50 (11.8%)		692 (13.8%)	658 (14.0%)	34 (11.5%)		284 (13.2%)	268 (13.3%)	16 (12.6%)	
N3	1,276 (17.8%)	1,203 (17.9%)	73 (17.3%)		875 (17.5%)	820 (17.4%)	55 (18.6%)		401 (18.7%)	383 (19.0%)	18 (14.2%)	
ER status				<0.001				<0.001				0.002
Negative	1,854 (25.9%)	1,691 (25.1%)	163 (38.6%)		1,253 (25.0%)	1,141 (24.2%)	112 (38.0%)		601 (28.0%)	550 (27.2%)	51 (40.2%)	
Positive	5,300 (74.1%)	5,041 (74.9%)	259 (61.4%)		3,754 (75.0%)	3,571 (75.8%)	183 (62.0%)		1,546 (72.0%)	1,470 (72.8%)	76 (59.8%)	
PR status				<0.001				<0.001				<0.001
Negative	2,861 (40.0%)	2,629 (39.1%)	232 (55.0%)		1,972 (39.4%)	1,816 (38.5%)	156 (52.9%)		889 (41.4%)	813 (40.2%)	76 (59.8%)	
Positive	4,293 (60.0%)	4,103 (60.9%)	190 (45.0%)		3,035 (60.6%)	2,896 (61.5%)	139 (47.1%)		1,258 (58.6%)	1,207 (59.8%)	51 (40.2%)	
HER2 status				0.044				0.056				0.515
Negative	5,178 (72.4%)	4,891 (72.7%)	287 (68.0%)		3,628 (72.5%)	3,429 (72.8%)	199 (67.5%)		1,550 (72.2%)	1,462 (72.4%)	88 (69.3%)	
Positive	1,976 (27.6%)	1,841 (27.3%)	135 (32.0%)		1,379 (27.5%)	1,283 (27.2%)	96 (32.5%)		597 (27.8%)	558 (27.6%)	39 (30.7%)	
Subtype				<0.001				<0.001				0.001
HR+, HER2–	4,126 (57.7%)	3,942 (58.6%)	184 (43.6%)		2,927 (58.5%)	2,795 (59.3%)	132 (44.7%)		1,199 (55.8%)	1,147 (56.8%)	52 (40.9%)	
HR+, HER2+	1,289 (18.0%)	1,207 (17.9%)	82 (19.4%)		897 (17.9%)	840 (17.8%)	57 (19.3%)		392 (18.3%)	367 (18.2%)	25 (19.7%)	
HR–, HER2+	687 (9.60%)	634 (9.42%)	53 (12.6%)		482 (9.63%)	443 (9.40%)	39 (13.2%)		205 (9.55%)	191 (9.46%)	14 (11.0%)	
HR–, HER2–	1,052 (14.7%)	949 (14.1%)	103 (24.4%)		701 (14.0%)	634 (13.5%)	67 (22.7%)		351 (16.3%)	315 (15.6%)	36 (28.3%)	

Table 1 *(continued)*

Table 1 *(continued)*

Variables	Whole cohort				Training set				Validation set			
	All, N=7,154	Non-brain metastasis, N=6,732	Brain metastasis, N=422	P value	All, N=5,007	Non-brain metastasis, N=4,712	Brain metastasis, N=295	P value	All, N=2,147	Non-brain metastasis, N=2,020	Brain metastasis, N=127	P value
Bone metastasis				0.957				1.000				0.947
No	2,560 (35.8%)	2,410 (35.8%)	150 (35.5%)		1,751 (35.0%)	1648 (35.0%)	103 (34.9%)		809 (37.7%)	762 (37.7%)	47 (37.0%)	
Yes	4,594 (64.2%)	4,322 (64.2%)	272 (64.5%)		3,256 (65.0%)	3,064 (65.0%)	192 (65.1%)		1,338 (62.3%)	1,258 (62.3%)	80 (63.0%)	
Liver metastasis				0.009				0.004				0.781
No	5,309 (74.2%)	5,019 (74.6%)	290 (68.7%)		3,706 (74.0%)	3,509 (74.5%)	197 (66.8%)		1,603 (74.7%)	1,510 (74.8%)	93 (73.2%)	
Yes	1,845 (25.8%)	1,713 (25.4%)	132 (31.3%)		1,301 (26.0%)	1,203 (25.5%)	98 (33.2%)		544 (25.3%)	510 (25.2%)	34 (26.8%)	
Lung metastasis				<0.001				<0.001				<0.001
No	5,032 (70.3%)	4,800 (71.3%)	232 (55.0%)		3,529 (70.5%)	3,365 (71.4%)	164 (55.6%)		1,503 (70.0%)	1,435 (71.0%)	68 (53.5%)	
Yes	2,122 (29.7%)	1,932 (28.7%)	190 (45.0%)		1,478 (29.5%)	1,347 (28.6%)	131 (44.4%)		644 (30.0%)	585 (29.0%)	59 (46.5%)	

Central: code C500 and C501; Inner: code C502 and C503; Outer: code C504, C505 and C506; Overlap: code C508; Unknown: code C509. From SEER program coding and staging manual 2021. IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; ER, estrogen receptor; PR, progesterone receptor; HR, hormone receptor; HER2, human epidermal growth factor receptor 2.

Table 2 Multivariate logistic regression analysis of predictive factors of brain metastasis in training set

Variables	OR	95% CI	P value
Age			0.025
<45	Reference	Reference	
45–64	1.541	1.064–2.230	0.022
>64	1.178	0.790–1.759	0.422
Grade			0.964
I	Reference	Reference	
II	1.157	0.647–2.068	0.622
III	1.172	0.647–2.124	0.601
IV	1.167	0.244–5.582	0.846
Histological type			0.513
IDC	Reference	Reference	
ILC	0.850	0.500–1.443	0.546
Others	0.737	0.414–1.312	0.300
Tumor size (cm)			0.081
≤2	Reference	Reference	
>2 and ≤5	0.694	0.491–0.982	0.039
>5	0.860	0.607–1.217	0.394
Subtype			0.001
HR+, HER2–	Reference	Reference	
HR+, HER2+	1.286	0.922–1.793	0.139
HR–, HER2+	1.558	1.050–2.313	0.028
HR–, HER2–	1.921	1.376–2.681	<0.001
Liver metastasis			
No	Reference	Reference	
Yes	1.280	0.988–1.658	0.061
Lung metastasis			
No	Reference	Reference	
Yes	1.855	1.450–2.375	<0.001

OR, odds ratio; CI, confidence interval; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; HR, hormone receptor; HER2, human epidermal growth factor receptor 2.

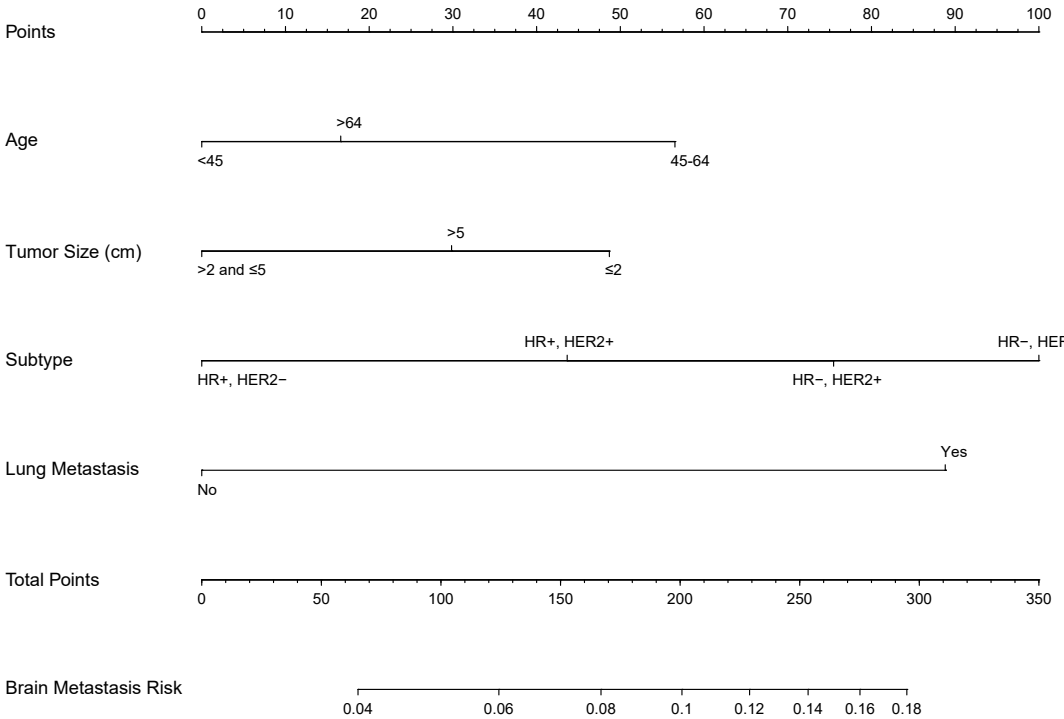

Figure 1 A nomogram for predicting brain metastasis in patients with de novo stage IV breast cancer.

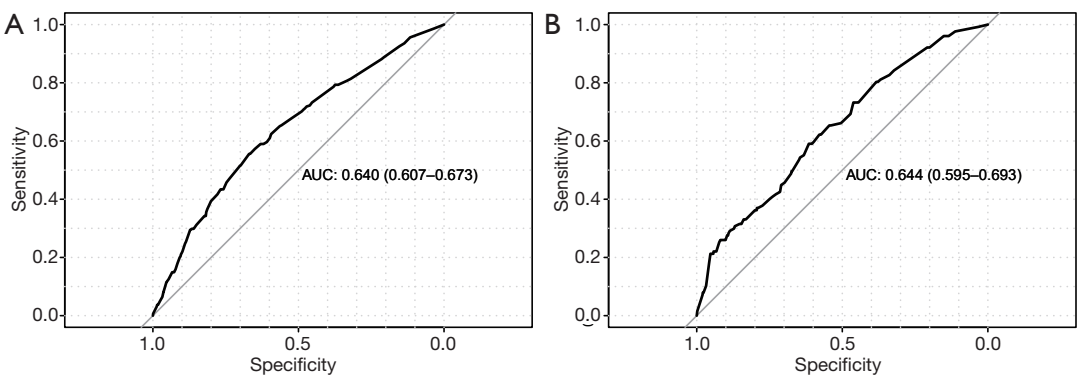

Figure 2 Receiver operating characteristic (ROC) curves of the training set and the validation set. (A) The nomogram had an area under curve (AUC) of 0.640 [95% confidence interval (CI): 0.607 to 0.673] in the training set. (B) The nomogram had an AUC of 0.644 (95% CI: 0.595 to 0.693) in the validation set.

Nomogram performance

The nomogram performance at predicting BM was evaluated according to discrimination, calibration, and clinical utility. The discrimination was evaluated with the ROC curves of the training set (*Figure 2A*) and the validation set (*Figure 2B*). The nomogram had discriminatory ability with an area under curve (AUC) of 0.640 [95% confidence interval (CI): 0.607 to 0.673] in the training set, and 0.644 (95% CI: 0.595 to 0.693) in the validation set.

Then, we performed the calibration of the nomogram internally by 1,000 bootstrap resamples with a calibration plot in the training set (*Figure 3*). Both the bias-corrected curve and the apparent curve were close to the ideal curve, demonstrating that the nomogram fitted well internally.

To evaluate the clinical utility of the nomogram,

we defined the predictive risk of BM as a new variable and performed Cox regression analysis together with other confirmed independent prognostic factors of breast cancer, including the degree of liver involvement, surgery, chemotherapy, and so on (*Table 3*). Multivariate cox regression analysis showed that the BM prediction risk obtained from the nomogram was one of the most significant variables in predicting OS [P<0.001, hazard ratio (HR): 1.499, 95% CI: 1.403 to 1.602].

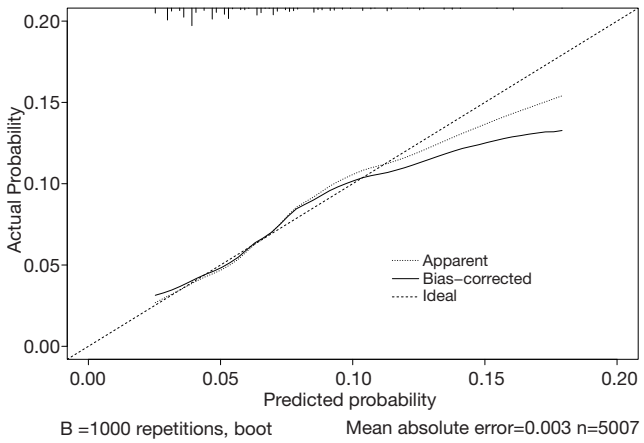

Figure 3 Calibration curves of the nomogram for brain metastasis (bootstrap =1,000 repetitions).

We then grouped participants according to the predicted risk of BM and observed significant differences between the two groups on the Kaplan-Meier curves (*Figure 4*, P<0.001). The results suggested that this predictive risk of BM could act as a prognostic indicator, participants with a higher predictive risk of BM had worse prognoses.

The clinical utility of this nomogram was also demonstrated by conducting a virtual trial in the validation set to prevent BM. Our nomogram could be used to identify a small subset of patients at high risk of BM to receive prophylactic intervention. We used the nomogram to predict individual risk of BM and set different thresholds as intervention conditions. *Table 4* shows how we evaluated the health economic value of our model in preventing BM by estimating the number of participants who needed preventive intervention and the number of BM successfully prevented. If all patients were given prophylactic treatment without selection, although the prevention of BM could have been maximized, many patients would have been overtreated. If we only included patients with a predictive risk of BM >5% in our nomogram, only 54.3% (1,165/2,147) of the population would be treated, but 70.9% (90/127) of

Table 3 Univariate and multivariate Cox regression analysis of prognostic factors for overall survival in the whole cohort

Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Race			<0.001			<0.001
Black	Reference	Reference		Reference	Reference	
White	0.758	0.700–0.821	<0.001	0.822	0.758–0.892	<0.001
Others	0.691	0.605–0.789	<0.001	0.763	0.668–0.873	<0.001
Marital status						
Unmarried	Reference	Reference		Reference	Reference	
Married	0.697	0.655–0.743	<0.001	0.755	0.708–0.806	<0.001
Primary site			<0.001			0.001
Central	Reference	Reference		Reference	Reference	
Inner	0.942	0.811–1.094	0.431	0.946	0.814–1.099	0.468
Outer	1.014	0.891–1.154	0.837	1.018	0.894–1.159	0.789
Overlap	0.997	0.871–1.142	0.968	0.977	0.852–1.119	0.733
Unknown	1.202	1.051–1.373	0.007	1.152	1.007–1.318	0.039

Table 3 *(continued)*

Table 3 *(continued)*

Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Grade			<0.001			<0.001
I	Reference	Reference		Reference	Reference	
II	1.240	1.078–1.426	0.003	1.325	1.150–1.528	<0.001
III	1.701	1.483–1.950	<0.001	1.978	1.712–2.284	<0.001
IV	1.962	1.333–2.889	0.001	2.359	1.597–3.485	<0.001
Histological type			0.021			<0.001
IDC	Reference	Reference		Reference	Reference	
ILC	1.001	0.898–1.116	0.979	1.317	1.173–1.478	<0.001
Others	0.834	0.733–0.949	0.006	1.029	0.903–1.172	0.666
N stage			0.001			<0.001
N0	Reference	Reference		Reference	Reference	
N1	0.868	0.801–0.941	0.001	0.824	0.760–0.894	<0.001
N2	0.826	0.742–0.919	<0.001	0.888	0.796–0.991	0.034
N3	0.873	0.792–0.964	0.007	0.942	0.851–1.042	0.244
Liver metastasis						
No	Reference	Reference		Reference	Reference	
Yes	1.580	1.477–1.691	<0.001	1.514	1.410–1.625	<0.001
Surgery of primary site			<0.001			<0.001
No	Reference	Reference		Reference	Reference	
Yes	0.557	0.521–0.595	<0.001	0.590	0.550–0.633	<0.001
Unknown	0.977	0.629–1.517	0.916	1.027	0.658–1.602	0.908
Surgery of distant site			0.001			0.064
No	Reference	Reference		Reference	Reference	
Yes	0.758	0.656–0.875	<0.001	0.841	0.727–0.972	0.019
Unknown	1.272	0.477–3.391	0.631	1.079	0.400–2.910	0.881
Chemotherapy						
No/unknown	Reference	Reference		Reference	Reference	
Yes	0.690	0.648–0.735	<0.001	0.588	0.549–0.629	<0.001
Radiotherapy						
No/unknown	Reference	Reference		Reference	Reference	
Yes	0.817	0.765–0.872	<0.001	1.005	0.940–1.075	0.885
Brain metastasis prediction						
Below mean	Reference	Reference		Reference	Reference	
Above mean	1.522	1.430–1.621	<0.001	1.499	1.403–1.602	<0.001

HR, hazard ratio; CI, confidence interval; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma.

subsequent BM would be potentially prevented, with the actual BM rate decreasing to 46.5% of the original, with the assumption that prophylactic intervention could reduce the risk of BM by 75%. *Table 4* displays several thresholds set for prophylactic treatment and the corresponding effects on the prevention of BM.

Discussion

BM is associated with declined quality of life caused by progressive neurologic impairment, which has been an increasingly serious problem in the treatment of breast cancer. In today's clinical practice, clinicians and researchers are increasingly interested in predictive models designed to predict the occurrence of clinical events or therapeutic effects. Nomogram is such a tool to predict personalized risk for patients and provide a basis for clinical decision-making. In our study, we summarized the specific clinicopathological characteristics of BM and developed a nomogram to predict BM in *de novo* stage IV breast cancer patients, thus helping clinicians identify groups at high-risk of BM and enabling the undertaking of early preventive interventions to improve their prognoses.

We have drawn some interesting conclusions from this nomogram. The most significant predictors in the nomogram were age, tumor size, subtype, and the degree of lung involvement. So far, the relationship between age and BM has been inconclusive. Some studies have indicated that younger patients experience a higher risk of developing BM (31-33). On the contrary, other studies have shown that aging increases the risk of BM (30,34). Different from these prior studies, the relationship between age and BM in our study was neither a straightforward positive nor negative correlation. The results showed that patients aged 45–64 had the highest risk of BM, followed by patients over 64 years old, and the lowest risk was at under 45 years old. The small number of young patients in the SEER database is a possible reason for this difference of age correlation with BM. Thus, it is necessary to carry out studies on larger populations in the future, especially in China, where young patients with breast cancer account for a higher proportion of the broader population (35). Nevertheless, our results were still supported by a previous

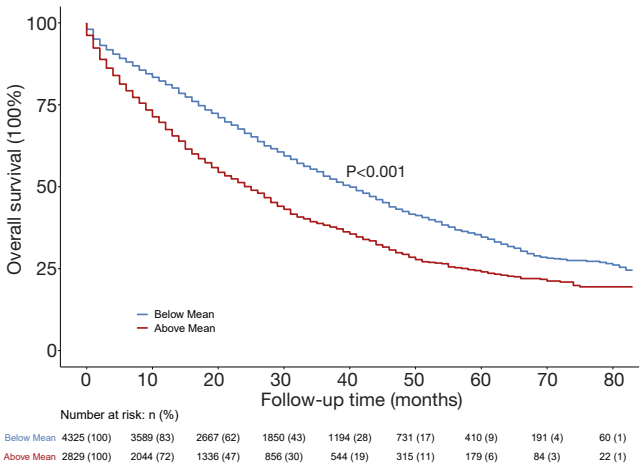

Figure 4 Kaplan-Meier curves comparing the overall survival for patients with the predictive risk of brain metastasis above or below mean.

Table 4 Clinical utility of the nomogram evaluated by the virtual trial with several thresholds for prophylactic treatment and the corresponding effects on the prevention of brain metastasis

Thresholds	Cases with PT	Potential BCBM cases with PT	Risk reduction of 50%		Risk reduction of 67%		Risk reduction of 75%	
			Cases prevented from BCBM	Cases of BCBM	Cases prevented from BCBM	Cases of BCBM	Cases prevented from BCBM	Cases of BCBM
0%	2,147	127 (100%)	64 (50.4%)	63 (49.6%)	85 (66.9%)	42 (33.1%)	95 (74.8%)	32 (25.2%)
3.6%	1,716	117 (92.1%)	59 (46.5%)	68 (53.5%)	78 (61.4%)	49 (38.6%)	88 (69.3%)	39 (30.7%)
4.3%	1,348	102 (80.3%)	51 (40.2%)	76 (59.8%)	68 (53.5%)	59 (46.5%)	77 (60.6%)	50 (39.4%)
5%	1,165	90 (70.9%)	45 (35.4%)	82 (64.6%)	60 (47.2%)	67 (52.8%)	68 (53.5%)	59 (46.5%)
6%	861	75 (59.1%)	38 (29.9%)	89 (70.1%)	50 (39.4%)	77 (60.6%)	56 (44.1%)	71 (55.9%)
8%	450	46 (36.2%)	23 (18.1%)	104 (81.9%)	31 (24.4%)	96 (75.6%)	35 (27.6%)	92 (72.4%)
100%	0	0 (0%)	0 (0%)	127 (100%)	0 (0%)	127 (100%)	0 (0%)	127 (100%)

PT, prophylactic therapy; BCBM, breast cancer brain metastasis.

large retrospective study which covered a population-based sample of 238,726 patients diagnosed with invasive breast cancer between 2010 and 2013 in US for whom the presence or absence of BM at diagnosis was known. Of the total of 238,726 patients, 5.68% of patients were aged 18–40, 41.13% were aged 41–60, 43.44% were aged 61–80, and 9.75% were older than 80. Consistent results were obtained in this study that patients between the ages of 41 and 60 had the highest risk of BM, followed by patients over the age of 60, and patients under the age of 41 (36). Tumor size was statistically an independent factor in our model, small tumors had the highest risk of BM, followed by large tumors, and intermediate-sized tumors had the lowest risk. This conclusion was controversial to common sense because many prior studies have demonstrated that progressively increasing tumor size was associated with an increased risk of BM (37,38). However, our conclusion was supported by the study of Tham *et al.*, in which the tumor size was closely related to BM, but the risk of BM did not always increase or decrease with tumor size (30). In our study, patients with TNBC had the highest risk of BM, followed by HR–/HER2+, HR+/HER2+, and HR+/HER2– had the lowest risk. The association between subtype and BM has been extensively discussed in previous studies (31,32,36,38,39), which was consistent with the conclusion of our study. Besides, the results of our study showed that patients with lung metastasis had a higher risk of BM. However, patients with liver or bone metastasis were not associated with a significant increase in the probability of BM. Slimane *et al.* also drew the same conclusion in their study (29). The mechanisms and clinical implications behind this phenomenon require further investigation. As for the pathological type of breast cancer, among the existing studies on the risk factors of BM, only two studies have reported that pathological type is related to BM. Similar conclusions were drawn from these two retrospective studies, suggesting that IDC histology is associated with higher risk for BM (30,40). Consistent with most relative studies, our study has not found a significant relationship between pathological type and BM, so we believe that pathological type is not related to BM.

Our nomogram has the following clinical application value by the identification of BM high-risk patients. First, clinicians can screen brain MRI regularly for high-risk patients to achieve early diagnosis and treatment. Second, in the aspect of systematic treatment, appropriate treatment lines of anticancer drugs can be moved forward for high-risk patients to prevent BM, especially in HER2-positive breast cancer, due to the excellent efficacy of novel targeted drugs. Third, in the aspect of local treatment, prophylactic cranial irradiation (PCI) is currently recommended for patients with small-cell lung cancer due to the significant rate of occult BM, resulting in reduced incidence of BM and improved survival outcomes (41-43). Although no guidelines recommended PCI for breast cancer to prevent BM currently, future randomized clinical trials of PCI for BM prevention might be undertaken in selected patients since high-risk groups can be identified by our nomogram. In our study, the health economic value of this nomogram in predicting BM has already been demonstrated by estimating the number of patients who needed preventive intervention and the number of patients whose BM were successfully prevented.

Our study had some major strengths and prominent observations. First, this was the first exploration of establishing a prediction model of BM using the SEER database which included about 30% of the US population, the clinicopathological features of the patients with BM we described were highly generalized and may better reflect population experience than previous studies limited to data from a single cancer center. Second, previous studies using clinicopathological features of the whole breast cancer population to predict BM had great limitations, because these features overlap with factors that increase the risk of distant metastasis, such as high histological grade and tumor stage. Our study attempted to overcome this by incorporating *de novo* MBC patients from the SEER database to balance and eliminate the influence of distant metastasis-related features, and to summarize the specific clinicopathological characteristics related to BM.

We acknowledge some notable limitations in our study. First of all, SEER data do not contain follow-up information about recurrence and metastasis of breast cancer, so we could only describe whether patients had BM at the time of initial diagnosis, but could not include patients who developed BM at the later stage of the disease course. Second, since the guidelines do not recommend routine BM screening for patients without CNS symptoms, we might well have underestimated the actual BM rate in newly diagnosed breast cancer. Third, the probability provided by our nomogram using a logistic regression model was binary and not time-related because we only knew the metastasis status at diagnosis. Fourth, the performance of our predictive model (with an AUC of around 0.64) was reasonable, but not great. This was due to a lack of information in the SEER database, such as disease

recurrence, subsequent sites of disease involvement, and the treatment before BM. In view of the fact that this was the first exploration of establishing a prediction model of BM using the SEER database, our nomogram would represent a good compromise. Fifth, since our study was a retrospective study, the conclusions raised still need to be further verified in prospective studies with a larger amount of data.

Acknowledgments

We gratefully acknowledge the efforts of the Surveillance, Epidemiology, and End Results (SEER) registries for the establishment and management of the database.
Funding: This study was supported by National Key R&D Program of China (2016YFC0901302) and Interdisciplinary Clinical Research Project of Peking University First Hospital (2019CR38).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at http://dx.doi.org/10.21037/atm-21-1808

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-21-1808). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All the procedures followed were in accordance with the Helsinki Declaration of the World Medical Association (as revised in 2013). The data released by the SEER database was publicly available and therefore did not require informed patient consent.

References

1. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391:1023-75.
2. Arvold ND, Oh KS, Niemierko A, et al. Brain metastases after breast-conserving therapy and systemic therapy: incidence and characteristics by biologic subtype. Breast Cancer Res Treat 2012;136:153-60.
3. Barnholtz-Sloan JS, Sloan AE, Davis FG, et al. Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System. J Clin Oncol 2004;22:2865-72.
4. Frisk G, Svensson T, Bäcklund LM, et al. Incidence and time trends of brain metastases admissions among breast cancer patients in Sweden. Br J Cancer 2012;106:1850-3.
5. Smedby KE, Brandt L, Bäcklund ML, et al. Brain metastases admissions in Sweden between 1987 and 2006. Br J Cancer 2009;101:1919-24.
6. Lin NU, Bellon JR, Winer EP. CNS metastases in breast cancer. J Clin Oncol 2004;22:3608-17.
7. Klos KJ, O'Neill BP. Brain metastases. Neurologist 2004;10:31-46.
8. Lin NU, Borges V, Anders C, et al. Intracranial Efficacy and Survival With Tucatinib Plus Trastuzumab and Capecitabine for Previously Treated HER2-Positive Breast Cancer With Brain Metastases in the HER2CLIMB Trial. J Clin Oncol 2020;38:2610-9.
9. Murthy RK, Loi S, Okines A, et al. Tucatinib, Trastuzumab, and Capecitabine for HER2-Positive Metastatic Breast Cancer. N Engl J Med 2020;382:597-609.
10. Awada A, Colomer R, Inoue K, et al. Neratinib Plus Paclitaxel vs Trastuzumab Plus Paclitaxel in Previously Untreated Metastatic ERBB2-Positive Breast Cancer: The NEfERT-T Randomized Clinical Trial. JAMA Oncol 2016;2:1557-64.
11. Freedman RA, Gelman RS, Anders CK, et al. TBCRC 022: A Phase II Trial of Neratinib and Capecitabine for Patients With Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer and Brain Metastases. J Clin Oncol 2019;37:1081-9.
12. Saura C, Oliveira M, Feng YH, et al. Neratinib Plus Capecitabine Versus Lapatinib Plus Capecitabine in HER2-Positive Metastatic Breast Cancer Previously Treated With ≥ 2 HER2-Directed Regimens: Phase III

NALA Trial. J Clin Oncol 2020;38:3138-49.
13. Krop IE, Lin NU, Blackwell K, et al. Trastuzumab emtansine (T-DM1) versus lapatinib plus capecitabine in patients with HER2-positive metastatic breast cancer and central nervous system metastases: a retrospective, exploratory analysis in EMILIA. Ann Oncol 2015;26:113-9.
14. Montemurro F, Delaloge S, Barrios CH, et al. Trastuzumab emtansine (T-DM1) in patients with HER2-positive metastatic breast cancer and brain metastases: exploratory final analysis of cohort 1 from KAMILLA, a single-arm phase IIIb clinical trial. Ann Oncol 2020;31:1350-8.
15. Modi S, Park H, Murthy RK, et al. Antitumor Activity and Safety of Trastuzumab Deruxtecan in Patients With HER2-Low-Expressing Advanced Breast Cancer: Results From a Phase Ib Study. J Clin Oncol 2020;38:1887-96.
16. Modi S, Saura C, Yamashita T, et al. Trastuzumab Deruxtecan in Previously Treated HER2-Positive Breast Cancer. N Engl J Med 2020;382:610-21.
17. Palma DA, Olson R, Harrow S, et al. Stereotactic ablative radiotherapy versus standard of care palliative treatment in patients with oligometastatic cancers (SABR-COMET): a randomised, phase 2, open-label trial. Lancet 2019;393:2051-8.
18. Andrews DW, Scott CB, Sperduto PW, et al. Whole brain radiation therapy with or without stereotactic radiosurgery boost for patients with one to three brain metastases: phase III results of the RTOG 9508 randomised trial. Lancet 2004;363:1665-72.
19. Gradishar WJ, Anderson BO, Abraham J, et al. Breast Cancer, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network 2020;18:452-78.
20. Cardoso F, Senkus E, Costa A, et al. 4th ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 4). Ann Oncol 2018;29:1634-57.
21. Ramakrishna N, Temin S, Chandarlapaty S, et al. Recommendations on Disease Management for Patients With Advanced Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer and Brain Metastases: ASCO Clinical Practice Guideline Update. J Clin Oncol 2018;36:2804-7.
22. Schade B, Rao T, Dourdin N, et al. PTEN deficiency in a luminal ErbB-2 mouse model results in dramatic acceleration of mammary tumorigenesis and metastasis. J Biol Chem 2009;284:19018-26.
23. Pestalozzi BC, Zahrieh D, Price KN, et al. Identifying breast cancer patients at risk for Central Nervous System (CNS) metastases in trials of the International Breast Cancer Study Group (IBCSG). Ann Oncol 2006;17:935-44.
24. Lin NU, Claus E, Sohl J, et al. Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases. Cancer 2008;113:2638-45.
25. Yerushalmi R, Woods R, Kennecke H, et al. Patterns of relapse in breast cancer: changes over time. Breast Cancer Res Treat 2010;120:753-9.
26. Albiges L, André F, Balleyguier C, et al. Spectrum of breast cancer metastasis in BRCA1 mutation carriers: highly increased incidence of brain metastases. Ann Oncol 2005;16:1846-7.
27. Song Y, Barry WT, Seah DS, et al. Patterns of recurrence and metastasis in BRCA1/BRCA2-associated breast cancers. Cancer 2020;126:271-80.
28. Evans AJ, James JJ, Cornford EJ, et al. Brain metastases from breast cancer: identification of a high-risk group. Clin Oncol (R Coll Radiol) 2004;16:345-9.
29. Slimane K, Andre F, Delaloge S, et al. Risk factors for brain relapse in patients with metastatic breast cancer. Ann Oncol 2004;15:1640-4.
30. Tham YL, Sexton K, Kramer R, et al. Primary breast cancer phenotypes associated with propensity for central nervous system metastases. Cancer 2006;107:696-704.
31. Aversa C, Rossi V, Geuna E, et al. Metastatic breast cancer subtypes and central nervous system metastases. Breast 2014;23:623-8.
32. Hung MH, Liu CY, Shiau CY, et al. Effect of age and biological subtype on the risk and timing of brain metastasis in breast cancer patients. PLoS One 2014;9:e89389.
33. Xue J, Peng G, Yang JS, et al. Predictive factors of brain metastasis in patients with breast cancer. Med Oncol 2013;30:337.
34. Graesslin O, Abdulkarim BS, Coutant C, et al. Nomogram to predict subsequent brain metastasis in patients with metastatic breast cancer. J Clin Oncol 2010;28:2032-7.
35. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. Lancet Oncol 2014;15:e279-89.
36. Martin AM, Cagney DN, Catalano PJ, et al. Brain Metastases in Newly Diagnosed Breast Cancer: A Population-Based Study. JAMA Oncol 2017;3:1069-77.
37. Azim HA, Abdel-Malek R, Kassem L. Predicting Brain Metastasis in Breast Cancer Patients: Stage Versus Biology. Clin Breast Cancer 2018;18:e187-95.
38. Kim YJ, Kim JS, Kim IA. Molecular subtype predicts incidence and prognosis of brain metastasis from breast cancer in SEER database. J Cancer Res Clin Oncol

2018;144:1803-16.
39. Genre L, Roché H, Varela L, et al. External validation of a published nomogram for prediction of brain metastasis in patients with extra-cerebral metastatic breast cancer and risk regression analysis. Eur J Cancer 2017;72:200-9.
40. Saip P, Cicin I, Eralp Y, et al. Identification of patients who may benefit from the prophylactic cranial radiotherapy among breast cancer patients with brain metastasis. J Neurooncol 2009;93:243-51.
41. Al Feghali KA, Ballout RA, Khamis AM, et al. Prophylactic Cranial Irradiation in Patients With Non-Small-Cell Lung Cancer: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Front Oncol 2018;8:115.
42. Maeng CH, Song JU, Shim SR, et al. The Role of Prophylactic Cranial Irradiation in Patients With Extensive Stage Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. J Thorac Oncol 2018;13:840-8.
43. Yin X, Yan D, Qiu M, et al. Prophylactic cranial irradiation in small cell lung cancer: a systematic review and meta-analysis. BMC Cancer 2019;19:95.

(English Language Editor: J. Jones)

Cite this article as: Sun MS, Liu YH, Ye JM, Liu Q, Cheng YJ, Xin L, Xu L. A nomogram for predicting brain metastasis in patients with de novo stage IV breast cancer. Ann Transl Med 2021;9(10):853. doi: 10.21037/atm-21-1808

