Open Access Full Text Article

ORIGINAL RESEARCH

Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: a population-based study

Weikai Xiao[1,*]
Shaoquan Zheng[1,*]
Anli Yang[2,*]
Xingcai Zhang[3,*]
Yutian Zou[1]
Hailin Tang[1]
Xiaoming Xie[1]

[1]Department of Breast Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, People's Republic of China; [2]Department of Medicine; Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA; [3] Department of Applied Physics, Harvard John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA 02138, USA

*These authors contributed equally to this work

Correspondence: Hailin Tang; Xiaoming Xie
Department of Breast Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, 651 Dongfeng East Road, Guangzhou, 510060, People's Republic of China
Email tanghl@sysucc.org.cn; xiexm@sysucc.org.cn

Background: It was unclear whether breast cancer subtypes are associated with the risk of site-specific metastases. This study aimed to evaluate the relationship between molecular subtypes and distant metastatic sites and their prognostic significance.

Methods: We identified 295,213 patients with invasive breast cancer from 2010 to 2014 using the Surveillance, Epidemiology and End Results database. Subtypes were classified into four categories: hormone receptor (HR$^+$)/human epidermal growth factor receptor 2 (HER2$^-$), HR$^+$/HER2$^+$, HR$^-$/HER2$^+$, and triple-negative (HR$^-$/HER2$^-$). Logistic regression was used to assess the association between metastasis location and subtypes. Multivariate Cox models were used to estimate the overall survival (OS) of related factors.

Results: According to our study, 3.28%, 1.52%, 1.20%, and 0.35% of newly diagnosed breast cancers presented bone, lung, liver, and brain metastases at diagnosis, respectively. Both metastatic sites and subtypes significantly affected the OS after metastasis. In multivariate analysis, HR$^+$/HER2$^+$ subtype (OR as compared with HR$^+$/HER2$^-$ subtype, 1.30 [95% CI, 1.22–1.39]) significantly correlated with elevated bone metastasis risk, whereas HR$^-$/HER2$^+$ did not. Both HER2$^+$ subtypes (HR$^+$/HER2$^+$ and HR$^-$/HER2$^+$) were significantly associated with higher rates of liver, brain, and lung metastases, while the highest OR was observed in liver metastases. Triple-negative tumors had a higher rate of brain (OR, 1.95 [95% CI, 1.61–2.35]), liver (OR, 1.35 [95% CI, 1.20–1.51]), and lung metastases (OR, 1.34 [95% CI, 1.21–1.47]), but a significantly lower rate of bone metastases (OR, 0.64 [95% CI, 0.59–0.69]) than HR$^+$/HER2− tumors.

Conclusions: Breast cancer subtypes are associated with different metastatic patterns and confer different prognostic impacts. Molecular subtypes can identify patients at increased risk of site-specific metastases.

Keywords: breast cancer, molecular subtypes, metastasis behavior, prognosis, metastatic sites

Introduction

Despite the significant advances made in the treatment of breast cancer in the past few decades, the prognosis for most patients with distant metastases remains poor.[1–3] A deep understanding of the molecular phenotype of distant metastasis is crucial to pave the way for the earlier detection of metastases and more effective treatments.[4] The process by which cancer cells move from the primary site to distant organs through the bloodstream is not random.[5] Instead, some types of cancer cells undergo a series of molecular programs that preferentially target specific organs and nest there.[6–8] This behavior of finding a destination involves the escape of the cancer cells from the primary tumor (sometimes referred to as "seeds") through the interaction with the microenvironment (or "soil") to the target organ.[5]

Breast cancer subtypes (hormone receptor (HR⁺)/human epidermal growth factor receptor 2 (HER2⁻), HR⁺/HER2⁺, HR⁻/HER2⁺, and triple-negative [HR⁻/HER2⁻] subtypes) have been shown to have clinical utility in guiding therapeutic decisions.[4,9] The decision whether to perform endocrine or HER2-targeted therapies is based primarily on the molecular subtype status of the tumor. Some studies have demonstrated that the molecular subtypes of breast cancer are closely related to the different risks of early recurrence and metastasis, response to treatment, and overall survival (OS).[9,10]

The molecular characteristics of the primary tumor usually remain in the metastases.[9,11] It has been reported that the genetic characteristics of the brain,[11] lung,[12] and bone[13] metastases and the expression of HER2 and estrogen receptor (ER) in breast cancer are associated with increased risk of metastasis to specific organs. For example, subtypes of brain metastases are more likely to be HER2⁺ or triple-negative subtypes.[14] To date, very few studies have evaluated the effect of molecular subtypes on specific metastatic sites. In particular, the distribution of molecular subtypes at different metastatic sites has been poorly understood. A better understanding of the mode of metastatic disease may affect adjuvant treatment and monitoring decisions and determine which screening and follow-up strategies are appropriate once breast cancer has been diagnosed.[15]

The purpose of this study is to determine the association between molecular subtypes and metastatic patterns and their impact on prognosis based on a large real-world database, the Surveillance, Epidemiology and End Results (SEER) database.

Methods
Data collection
The patients' data were extracted from the SEER database,[16] which covers about 30% of the American population in 18 registries released in 2016. We obtained 312,035 patients diagnosed at 18 years old or above with primary and microscopically validated malignant breast cancer between January 1, 2010 and December 31, 2014 using SEER*Stat software (version 8.3.4). Furthermore, those with unknown distant metastatic status were excluded from this population and we further excluded the patients diagnosed by death certificate or biopsy without active follow-up and whose survival record was zero months, leaving 295,213 patients in the entire cohort and generating a study cohort of 12,972 patients harboring distant metastases at first diagnosis of breast cancer. Before initiating this study, we submitted a data use agreement to the SEER program and were officially granted access to the database. The case listing comprised information on the

following covariates: age, sex, race, marital status, insurance, molecular subtype, pathological grade, and distant metastatic sites. The Institutional Review Board of Sun Yat-sen University Cancer Center reviewed our study and waived the written consent forms because the patients could not be identified. We have obtained consents to publish this paper from all the participants of this study.

Statistical analysis
Patients were categorized as the following breast cancer molecular subtypes: HR⁺/HER2⁺, HR⁻/HER2⁺, HR⁺/HER2⁻, triple-negative (HR⁻/HER2⁻), and unknown subtypes. The incidence proportion was defined as the percentage of patients with distant metastases at first diagnosis of breast cancer among the breast cancer population in each subset and computed after stratification by age, sex, race, etc. The patients were divided by age at diagnosis: <40 years, 40–65 years, and >65 years. Races were classified as white, black, Asian or Pacific Islander, and American Indian/Alaska Native in accordance with the database record.

We used multivariable logistic regression to estimate the ORs for subtypes which might be predictive for having a specific kind of distant metastases at first diagnosis of breast cancer. This model was adjusted for all potential influential factors, including age, sex, race, marital status, insurance, AJCC7 T stage, AJCC7 N stage, subtypes, pathological types, and grade. Distant metastasis refers to the bone, liver, lung, and brain metastases, apart from other metastatic sites including lymph node metastases.

We defined OS as the time from the first diagnosis of breast cancer to death due to breast cancer or other causes. Kaplan–Meier estimates of OS were plotted by subtype and the extent of systemic disease.

We calculated 95% CIs for all OR*s* across the subsets. A *P*-value of 0.05 or less was determined as statistically significant. All *P*-values were two-tailed. Statistical analysis was conducted using SPSS statistical software (SPSS IBM Statistics 21; IBM Corporation, Armonk, NY, USA). Apart from the subtype distribution calculation, figures and tables were produced by Excel software (Microsoft 2010). Line charts of median survival and Kaplan–Meier curves were plotted by SAS software (version 9.4; SAS Institute Inc., Cary, NC, USA) and GraphPad Prism 7 (GraphPad Software, Inc., La Jolla, CA, USA).

Results
Of all patients diagnosed between 2010 and 2014, we identified 295,213 patients with invasive breast cancer from the

SEER database according to the inclusion criteria (Figure 1). The average age of diagnosis was 61.68 years (median: 61 years). The prevalence of invasive breast cancer was 68.1% in HR+/HER2− breast cancer patients (N=200,976), 9.53% in HR+/HER2+ breast cancer patients (N=28,150), 4.12% in HR−/HER2+ breast cancer patients (N=12,164), and 10.6% in triple-negative breast cancer patients (N=31,217) (Table 1). Around 3.28% of newly diagnosed breast cancer patients at diagnosis had bone metastases (N=9,688), 1.52% had lung metastases (N=4,482), 1.20% had liver metastases (N=3,553), and 0.35% had brain metastases (N=1,045) (Table 1). When patients were stratified by metastatic sites, there were statistical differences in age, race, marital status, insurance status, T-stage, N-stage, histology, tumor grade, and molecular subtype (all $P<0.05$) (Table 2).

Impact of molecular subtype on metastatic sites

Figure 2A shows subtype distributions in patients with bone, lung, liver, and brain metastases in breast cancer patients at diagnosis, respectively (Figure 2A). The percentage of HR+/HER2− subtype gradually decreased in patients presenting
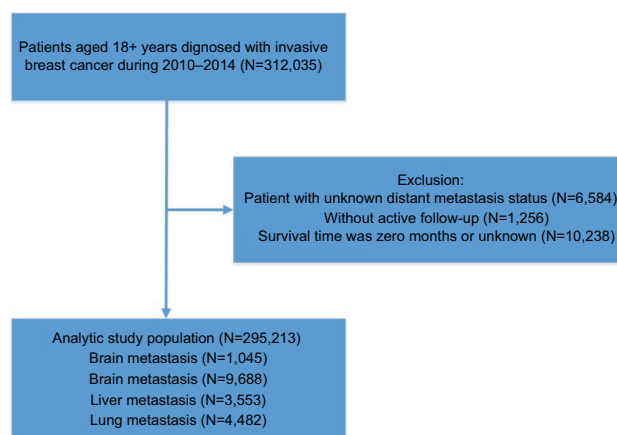
with bone (59.9%), lung (47.8%), liver (40.9%), and brain (38.8%) metastases compared with the cohort of all breast cancer patients (68.1%). The percentage of both HER2+ subtypes in the four metastatic sites increased, with the highest percentage in patients presenting with liver metastases (20.2% for HR+/HER2+, 13.1% for HR−/HER2+), compared with the entire cohort (9.53% for HR+/HER2+, 4.12% for HR−/HER2+). For the triple-negative subtype, the percentage of patients with bone metastases (8%) decreased compared to the entire cohort (10.6%). The proportion of visceral (lung, liver, and brain) metastases increased in the triple-negative subtypes, with the greatest increase in brain metastases (18.9%). In addition, this trend has been further verified in the patients with oligo-oran metastasis (Figure 2B). In addition, using a multivariable logistic regression model that adjusted for age, gender, race, insurance status, marital status, tumor size, and lymph node metastasis, it was found that patients with triple-negative and HER2+ subtypes had an increased risk of visceral (lung, liver, and brain) metastases, as shown in Table 3. Compared with HR+/HER2− tumors, both HER2+ subtypes were associated with significantly higher rates of liver, brain, and lung metastases, but their ORs gradually decreased in liver, brain, and lung metastases. Triple-negative tumors had a higher rate of brain (OR, 1.95 [95% CI, 1.61–2.35]), liver (OR, 1.35 [95% CI, 1.20–1.51]), and lung (OR, 1.34 [95% CI, 1.21–1.47]) metastases, but a significantly lower rate of bone metastases (OR, 0.64 [95% CI, 0.59–0.69]) than HR+/HER2− tumors. In addition, older age (>65 years) was associated with higher prevalence of lung metastases (OR, 1.75 [95% CI, 1.52–2.02]), but with lower liver metastases (OR, 0.69 [95% CI, 0.61–0.79]). Another interesting finding is that lobular cancer is associated with an increased risk of bone metastases (OR, 1.27 [95% CI, 1.2–1.35]), yet it is associated with a significant decrease in the risk of brain (OR, 0.65 [95% CI, 0.51–0.83]), lung (OR, 0.87 [95% CI, 0.77–0.97]), and liver (OR, 0.47 [95% CI, 0.42–0.54]) metastases.

Figure 1 Patient selection flow chart.

Table 1 The incidence of patients with distant metastases was stratified by subtypes

Subtypes	N (%)				
	Newly diagnosed breast cancer	Bone met	Lung met	Liver met	Brain met
HR+/HER2−	200,976 (68.1)	5,799 (59.9)	2,144 (47.8)	1,452 (40.9)	405 (38.8)
HR+/HER2+	28,150 (9.53)	1,319 (13.6)	598 (13.3)	717 (20.2)	164 (15.7)
HR−/HER2+	12,164 (4.12)	507 (5.23)	381 (8.50)	464 (13.1)	112 (10.7)
Triple-negative	31,217 (10.6)	775 (8.00)	712 (15.9)	474 (13.3)	195 (18.7)
Unknown	22,706 (7.69)	1,288 (13.3)	647 (14.4)	446 (12.6)	169 (16.2)
All subtypes	295,213 (100)	9,688 (100)	4,482 (100)	3,553 (100)	1,045 (100)

Abbreviations: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; +, denotes positive; −, denotes negative; met, metastasis; N, number.

Table 2 Clinicopathological and demographic characteristics of patients according to metastatic sites

Variables	N (%)				
	Brain met	**Liver met**	**Lung met**	**Bone met**	**All met**
Age at diagnosis (years)[a]	60 (5.7)	325 (9.1)	229 (5.1)	622 (6.4)	835 (6.4)
<40	60 (5.7)	325 (9.1)	229 (5.1)	622 (6.4)	835 (6.4)
40–65	675 (64.6)	2,116 (59.6)	2,340 (52.2)	5,431 (56.1)	7,123 (54.9)
>65	310 (29.7)	1,112 (31.3)	1913 (42.7)	3,635 (37.5)	5,014 (38.7)
Sex					
Female	1,034 (98.9)	3,530 (99.4)	4,409 (98.4)	9,574 (98.8)	12,826 (98.9)
Male	11 (1.1)	23 (0.6)	73 (1.6)	114 (1.2)	146 (1.1)
Race					
White	767 (73.4)	2,641 (74.3)	3,261 (72.8)	7,484 (77.3)	9,853 (76)
Black	207 (19.8)	636 (17.9)	851 (19)	1,498 (15.5)	2,168 (16.7)
Asian or Pacific Islander	63 (6.00)	241 (6.8)	316 (7.1)	612 (6.3)	819 (6.3)
American Indian/Alaska Native	6 (0.6)	22 (0.6)	34 (0.8)	54 (0.6)	82 (0.6)
Unknown	2 (0.2)	13 (0.4)	20 (0.4)	40 (0.4)	50 (0.4)
Marital status					
Married	430 (41.1)	1,561 (43.9)	1,745 (38.9)	4,202 (43.4)	5,548 (42.8)
Unmarried	557 (53.3)	1,780 (50.1)	2,464 (55)	4,959 (51.2)	6,706 (51.7)
Unknown	58 (5.6)	212 (6)	273 (6.1)	527 (5.4)	718 (5.5)
Insurance					
Insured	961 (92.0)	3,333 (93.8)	4,153 (92.7)	9,094 (93.9)	4,293 (33.1)
Uninsured	65 (6.20)	144 (4.1)	225 (5)	397 (4.1)	6,677 (51.5)
Unknown	19 (1.80)	76 (2.1)	104 (2.3)	197 (2)	2,002 (15.4)
T stage					
T0-2	290 (27.8)	1,155 (32.5)	1,259 (28.1)	3,207 (33.1)	12,189 (94)
T3-4	553 (52.9)	1,849 (52)	2,568 (57.3)	4,931 (50.9)	517 (4)
Unknown	202 (19.3)	549 (15.5)	655 (14.6)	1,550 (16)	266 (2.1)
N stage					
N0	277 (26.5)	808 (22.7)	998 (22.3)	2,441 (25.2)	3,273 (25.2)
N1	396 (37.9)	1,554 (43.7)	1,937 (43.2)	3,981 (41.1)	5,315 (41)
N2–3	246 (23.5)	806 (22.7)	1,057 (23.6)	2,232 (23)	3,027 (23.3)
Unknown	126 (12.1)	385 (10.8)	490 (10.9)	1,034 (10.7)	1,357 (10.5)
Histology					
Ductal	704 (67.4)	2,562 (72.1)	3,253 (72.6)	6,321 (65.2)	8,732 (67.3)
Lobular	77 (7.4)	360 (10.1)	296 (6.6)	1,743 (18)	1,936 (14.9)
Others	264 (25.3)	631 (17.8)	933 (20.8)	1,624 (16.8)	2,304 (17.8)
Grade					
1–2	326 (31.2)	1,175 (33.1)	1,536 (34.3)	4,214 (43.5)	5,144 (39.7)
3–4	427 (40.9)	1,591 (44.8)	1,910 (42.6)	3,108 (32.1)	4,764 (36.7)
Unknown	292 (27.9)	787 (22.2)	1,036 (23.1)	2,366 (24.4)	3,064 (23.6)
Subtypes					
HR$^+$/HER2$^-$	405 (38.8)	1,452 (40.9)	2,144 (47.8)	5,799 (59.9)	7,009 (54)
HR$^+$/HER2$^+$	164 (15.7)	717 (20.2)	598 (13.3)	1,319 (13.6)	1,826 (14.1)
HR$^-$/HER2$^+$	112 (10.7)	464 (13.1)	381 (8.5)	507 (5.2)	936 (7.2)
Triple-negative	195 (18.7)	474 (13.3)	712 (15.9)	775 (8)	1,450 (11.2)
Unknown	169 (16.2)	446 (12.6)	647 (14.4)	1,288 (13.3)	1,751 (13.5)

Notes: [a]Unknown age was removed from model owing to non-convergence.
Abbreviations: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; +, denotes positive; –, denotes negative; met, metastasis; N, number.

Impact of subtypes on survival

Multivariable Cox analysis of OS according to metastatic sites is summarized in Table 4. In all four kinds of distant metastases, HR$^+$/HER2$^+$ (vs HR$^+$/HER2$^-$) subtype is significantly associated with increased OS, while triple-negative subtype was significantly associated with a decreased OS.

Compared with HR$^+$/HER2$^-$, HR$^-$/HER2$^+$ subtype is significantly associated with worse prognosis in patients with bone metastases (HR, 1.27 [95% CI, 1.10–1.46]) and lung metastases (HR, 1.30 [95% CI, 1.11–1.52]), but not in patients with liver and brain metastases (P>0.05). The median survival of patients with specific metastatic status and active follow-up
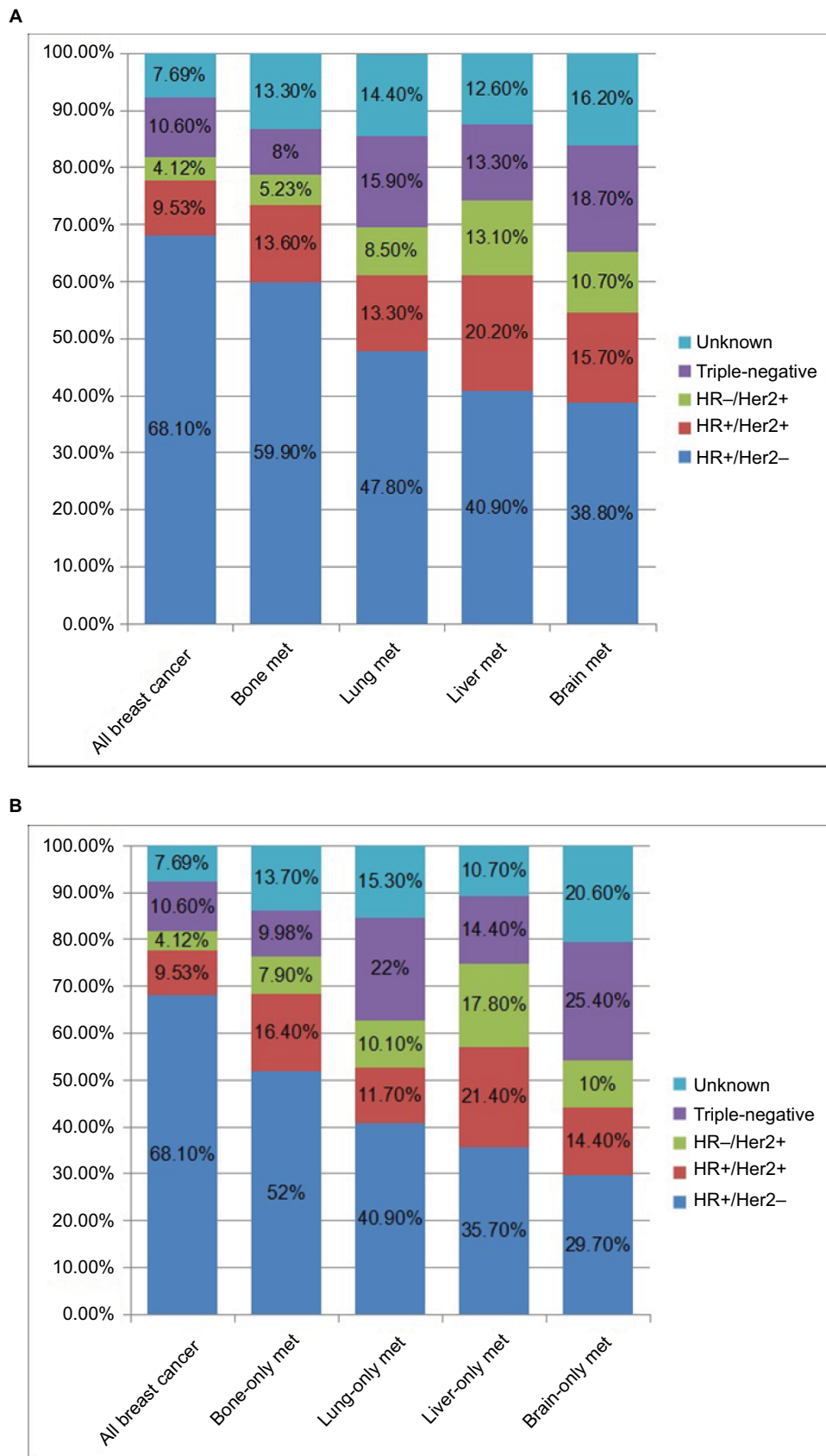
Figure 2 Distribution (percentage) of molecular subtypes in all breast cancer patients, and in patients with bone, lung, liver, and brain metastases (**A**); distribution (percentage) of molecular subtypes in all breast cancer patients and in patients with oligo-organ metastases (**B**).
Abbreviations: met, metastasis; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; +, denotes positive; –, denotes negative.

Table 3 Multivariate logistic regression for metastatic sites among newly diagnosed breast cancer[a]

Variables	OR (95% CI)			
	Brain met	**Liver met**	**Lung met**	**Brain met**
Age at diagnosis, years[b]				
<40	1 [Reference]	1[Reference]	1 [Reference]	1 [Reference]
40–65	1.39 (1.06–1.81)	0.80 (0.71–0.90)	1.29 (1.12–1.48)	NS
>65	NS	0.69 (0.61–0.79)	1.75 (1.52–2.02)	NS
Sex				
Female	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Male	NS	NS	1.76 (1.38–2.24)	1.30 (1.07–1.59)
Race				
White	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Black	1.23 (1.05–1.44)	1.19 (1.08–1.3)	1.28 (1.18–1.39)	1.11 (1.04–1.18)
Asian or Pacific Islander	0.76 (0.59–0.99)	0.83 (0.72–0.95)	NS	0.81 (0.74–0.88)
American Indian/Alaska Native	NS	NS	NS	NS
Unknown	NS	0.49 (0.28–0.85)	NS	0.53 (0.38–0.74)
Marital status				
Married	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Unmarried	1.46 (1.28–1.67)	1.34 (1.25–1.44)	1.51 (1.42–1.62)	1.41 (1.35–1.47)
Unknown	NS	NS	1.23 (1.07–1.41)	NS
Insurance				
Insured	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Uninsured	2.04 (1.57–2.65)	1.4 (1.18–1.68)	2.06 (1.78–2.38)	1.73 (1.54–1.94)
Unknown	0.61 (0.38–0.99)	0.74 (0.58–0.94)	NS	0.65 (0.55–0.76)
T stage				
T0-2	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
T3-4	2.17 (1.88–2.50)	1.85 (1.71–1.99)	2.36 (2.20–2.52)	1.81 (1.73–1.89)
Unknown	4.17 (3.30–5.27)	3.31 (2.88–3.79)	3.09 (2.72–3.51)	3.95 (3.62–4.3)
N stage				
N0	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
N1	3.45 (2.95–4.03)	4.73 (4.34–5.17)	5.37 (4.96–5.8)	4.79 (4.54–5.05)
N2–3	5.21 (4.36–6.22)	6.04 (5.45–6.68)	7.29 (6.65–7.98)	7.18 (6.75–7.63)
Unknown	5.01 (3.82–6.55)	8.35 (7.13–9.77)	8.55 (7.42–9.86)	8.51 (7.69–9.41)
Histology				
Ductal	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Lobular	0.65 (0.51–0.83)	0.87 (0.77–0.97)	0.47 (0.42–0.54)	1.27 (1.2–1.35)
Others	1.87 (1.57–2.23)	1.42 (1.28–1.59)	1.62 (1.48–1.77)	1.41 (1.32–1.52)
Grade				
1–2	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
3–4	1.34 (1.14–1.58)	1.50 (1.37–1.63)	1.50 (1.39–1.62)	1.19 (1.13–1.26)
Unknown	3.25 (2.68–3.96)	3.03 (2.72–3.39)	3.15 (2.86–3.48)	3.59 (3.36–3.83)
Subtypes				
HR[+]/HER2[−]	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
HR[+]/HER2[+]	2.04 (1.69–2.47)	2.52 (2.29–2.78)	1.41 (1.28–1.55)	1.3 (1.22–1.39)
HR[−]/HER2[+]	2.61 (2.09–3.26)	3.17 (2.83–3.56)	1.66 (1.47–1.86)	NS
Triple-negative	1.95 (1.61–2.35)	1.35 (1.20–1.51)	1.34 (1.21–1.47)	0.64 (0.59–0.69)
Unknown	1.26 (1.02–1.57)	1.11 (0.98–1.26)	1.07 (0.96–1.19)	0.83 (0.76–0.89)

Notes: [a]Only statistically significant results were presented ($P<0.05$); [b]Unknown age was removed from model owing to non-convergence.
Abbreviations: +, denotes positive; −, denotes negative; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; NS, not significant.

is stratified by subtypes and plotted in Figure 3. Figure 4 shows the Kaplan–Meier survival curve. Brain metastases harbor the worst outcomes, while bone metastases the best (11 months vs 30 months, respectively). The triple-negative subtype predicted the worst prognosis for all metastases, with a median survival (IQR) of 6.0 (2.0–13.0), 10.0 (4.0–19.0), 9.0 (3.0–17.0), and 11.0 (4.0–20.0) months in brain, bone, liver, and lung metastases, respectively. In contrast, HER2[+]/HR[+] cases show the best results among these metastatic sites, with an IQR of 25.0 (7.0–NR [not reached]), 41.0 (18.0–NR), 35.0 (12.0–NR), and 33.0 (15.0–NR) months, respectively.

Table 4 Multivariable Cox analysis of OS according to metastatic sites[a]

Variables	Hazard ratio (95% CI)			
	Brain met	**Liver met**	**Lung met**	**Bone met**
Age (years) at diagnosis				
<40	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
40–65	1.72 (1.15–2.57)	1.50 (1.23–1.82)	1.35 (1.09–1.68)	1.34 (1.16–1.56)
>65	2.76 (1.82–4.17)	2.60 (2.13–3.19)	2.02 (1.62–2.51)	2.06 (1.77–2.4)
Sex				
Female	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Male	NS	NS	NS	NS
Race				
White	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Black	NS	1.13 (1.00–1.27)	1.18 (1.06–1.31)	1.17 (1.08–1.27)
Asian or Pacific Islander	NS	NS	NS	NS
American Indian/Alaska Native	NS	NS	NS	NS
Unknown	NS	NS	NS	NS
Marital status				
Married	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Unmarried	1.22 (1.03–1.43)	1.30 (1.18–1.43)	1.34 (1.23–1.46)	1.36 (1.28–1.45)
Unknown	1.66 (1.18–2.32)	NS	NS	NS
Insurance				
Insured	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Uninsured	NS	1.44 (1.15–1.80)	NS	1.36 (1.17–1.57)
Unknown	NS	NS	NS	NS
T stage				
T0-2	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
T3-4	NS	1.13 (1.02–1.26)	NS	1.17 (1.09–1.25)
Unknown	NS	1.28 (1.1–1.48)	1.34 (1.17–1.54)	1.31 (1.19–1.44)
N stage				
N0	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
N1	NS	0.87 (0.78–0.98)	NS	NS
N2–3	NS	NS	NS	NS
Unknown	NS	NS	NS	1.16 (1.04–1.29)
Histology				
Ductal	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
Lobular	NS	NS	N S	NS
Others	NS	1.30 (1.13–1.49)	1.30 (1.17–1.46)	1.29 (1.18–1.41)
Grade				
1–2	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
3–4	NS	1.18 (1.05–1.32)	1.41 (1.27–1.56)	1.37 (1.27–1.48)
Unknown	NS	NS	1.19 (1.05–1.36)	1.14 (1.04–1.25)
Subtypes				
HR$^+$/HER2$^-$	1 [Reference]	1 [Reference]	1 [Reference]	1 [Reference]
HR$^+$/HER2$^+$	0.74 (0.57–0.95)	0.67 (0.59–0.78)	0.85 (0.73–0.98)	0.85 (0.77–0.94)
HR$^-$/HER2$^+$	NS	NS	1.30 (1.11–1.52)	1.27 (1.10–1.46)
Triple-negative	2.10 (1.68–2.62)	2.12 (1.86–2.42)	2.17 (1.93–2.44)	2.67 (2.42–2.95)
Unknown	1.61 (1.27–2.05)	1.21 (1.05–1.40)	1.56 (1.38–1.77)	1.39 (1.27–1.53)

Notes: [a]Only statistically significant results were presented (*P*<.05).
Abbreviations: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; NS, not significant; +, denotes positive; −, denotes negative.

Discussion

To our knowledge, this study is by far the largest comprehensive analysis of the relationship between molecular subtypes and distant metastasis patterns while controlling for multiple clinical-pathological variables. In this analysis of 295,213 patients, we confirmed the hypothesis that the risk of site-specific metastasis is subtype dependent. This large sample size allows us to more accurately estimate the prevalence in each metastatic site after stratification by molecular subtypes, and to assess differences in survival outcomes at different metastatic sites.

Overall, the most common distant metastasis was bone metastases (3.28%), followed by lung (1.52%), liver (1.20%),
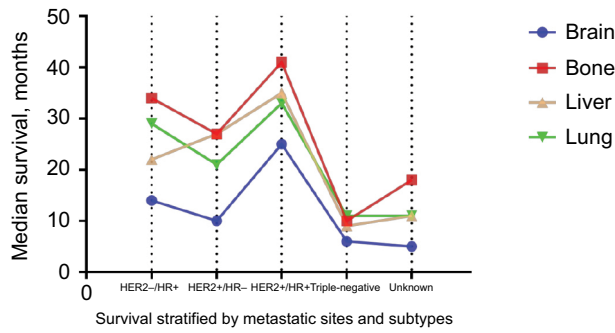
Figure 3 Median survival time (months) of different distant metastatic diseases stratified by molecular subtypes.

Abbreviations: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; +, denotes positive; −, denotes negative.

and brain metastasis (0.35%). In the multivariate analysis, compared with HR⁺/HER2⁻ tumors, both HER2⁺ subtypes were associated with significantly higher rates of liver, brain, and lung metastases, but the ORs gradually decreased in liver, brain, and lung metastases. Triple-negative tumors had a higher rate of the brain (OR, 1.95), liver (OR, 1.35), and lung metastases (OR, 1.34) but a significantly lower rate of bone metastases (OR, 0.64) than HR⁺/HER2⁻ tumors. The prognosis of specific metastatic site varied with subtypes, with HR⁺/HER2⁺ subtypes being the best and triple-negative subtypes being the worst.
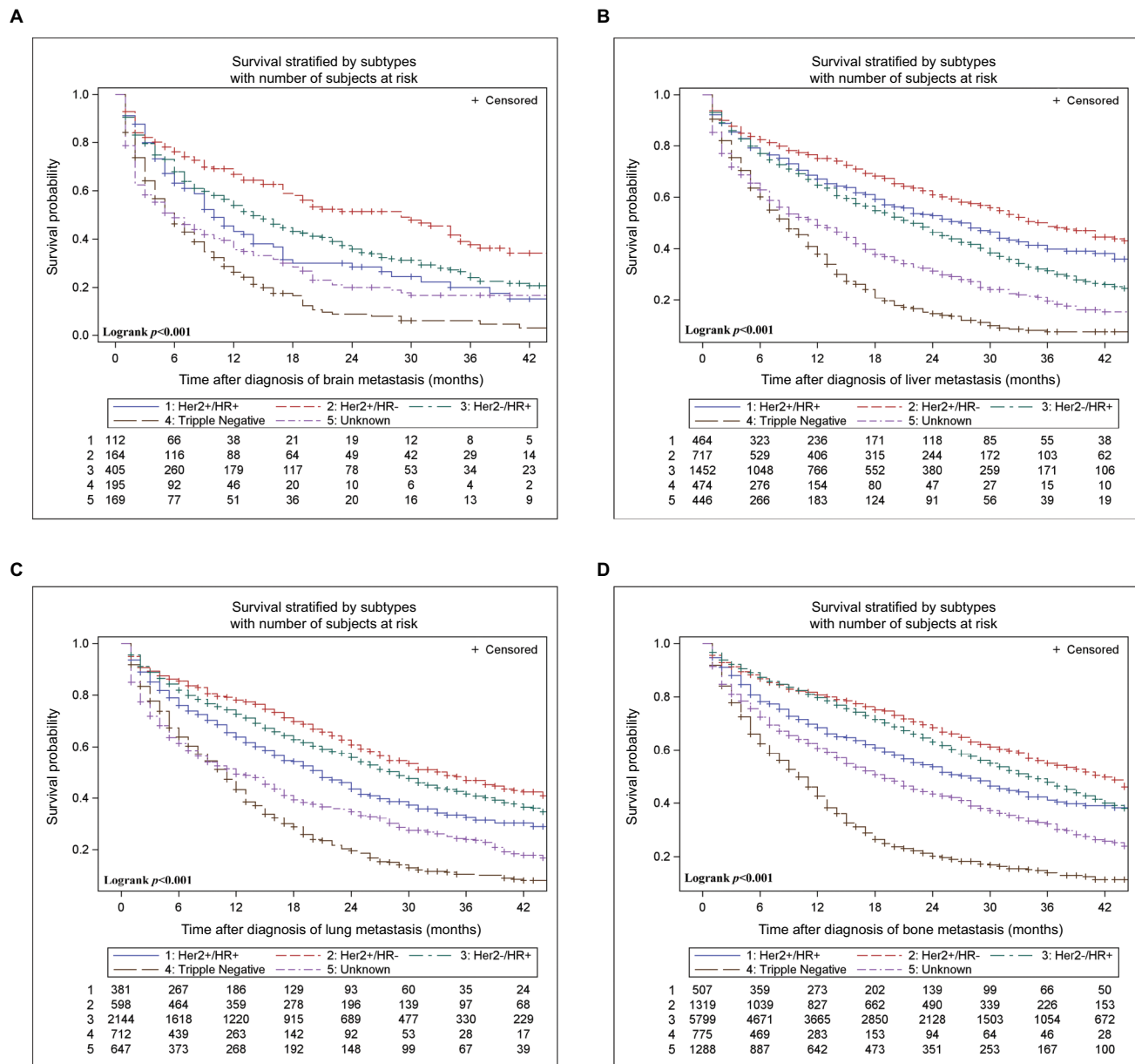
Figure 4 Kaplan–Meier survival curves and log-rank test results of different molecular subtypes in the brain (**A**), liver (**B**), lung (**C**), and bone (**D**) metastases.

Abbreviations: HER2, human epidermal growth factor receptor 2; HR, hormone receptor; +, denotes positive; −, denotes negative.

Previous studies have shown a significant difference in the time to distant recurrence.[9,17] ER⁻ tumors are associated with early recurrence, whereas ER⁺ tumors are associated with a more than 5-year sustained late risk.[18,19] Our study focused on distant metastases and found that patterns of metastasis were similar to that of early recurrence pattern of postoperative breast cancer. In our study, the risk of distant metastasis of HER2⁺ and triple-negative tumors was significantly increased. Previous studies have reported triple-negative breast cancer had a higher rate of brain metastasis in 6.2%–23.9% of patients;[9,20,21] a rate of 18.7% of triple-negative tumors was found among patients with brain metastasis in this study. In previous studies, ER⁻/HER2⁺ tumors had a higher risk of brain metastases than ER⁺/HER2⁻ tumors.[22–24] Our research also confirmed this point.

There are also substantial differences in baseline characteristics according to metastatic sites. For example, the impact of age on different metastatic sites is heterogeneous. Young breast cancer cases appear to have a higher risk of bone metastases, but the difference was not statistically different. The risk of liver metastases decreases with age, but the risk of lung metastases increases with age. For brain metastases, patients aged 40–65 years were at higher risk than other age groups. Unlike the study by Kennecke et al,[9] we found that larger tumor size (T3/4) was significantly associated with the incidence of distant metastases. The risk of bone metastases of lobular carcinoma is significantly higher than that of ductal carcinoma, but the risk of lung, liver, and brain metastases is reduced. The size of the primary tumor correlates with the prognosis of distant metastases, but there seems to be no significant correlation between regional lymph node metastasis and prognosis. These observations may reflect the impact of patient characteristics on the more aggressive disease tendencies.

This study confers several important implications for clinical practice. Firstly, breast cancer subtype can help identify patients at increased risk for specific metastatic sites. Secondly, the risk assessment of subtype-based site-specific metastasis can play an important role in the era of personalized medicine, and there is a need to work together to make advances in this area. In designing clinical trials, special attention should be paid to the matching of molecular subtypes and metastatic sites. A significant effect of molecular subtypes on distant metastases should also be considered during clinical examination or follow-up. Thirdly, our observations suggest that patients with bone metastases may have better survival outcomes, showing that these patients are more likely to benefit from primary tumor surgery.[25]

This study harbors several limitations. Firstly, the details of the subsequent metastasis during follow-up are not yet clear as SEER database does not provide this data. Secondly, we were unable to gather information on molecular subtypes of metastatic lesions, which may provide insights into the effects of molecular subtypes on specific metastatic sites.[26] Finally, we could not determine the exact location or number of each metastasis and the relevant information on treatment, so the impact of these factors on survival could not be obtained from our results.

Pre-emptive identification of patients at risk of distant metastases can improve the effectiveness of early diagnosis and interventions, and enhance the monitoring of the preferred location for metastasis in breast cancer patients. However, it was unclear whether breast cancer subtypes are associated with the risk of site-specific metastases. In this study, we apply by far the largest comprehensive analysis of the relationship between molecular subtypes and distant metastasis patterns while controlling for multiple clinical-pathological variables. In this analysis of 295,213 patients, we confirmed the hypothesis that the risk of site-specific metastasis is subtype dependent. This large sample size allows us to more accurately estimate the prevalence in each metastatic site after stratification by molecular subtypes, and to assess differences in survival outcomes at different metastatic sites. Our study can be further validated to assist clinical trial design and treatment and is of special importance for identifying patients at increased risk of site-specific metastases.

Conclusion

Breast cancer subtypes are associated with different metastatic patterns and confer different prognostic impacts. Molecular subtypes can identify patients at increased risk of site-specific metastases.

Data sharing statement

The datasets generated and/or analyzed during the current study are available in the SEER Program [https://seer.cancer.gov/].

Acknowledgment

Author contributions

Study concept and design: WX, SZ, AY, HT, XX; acquisition, analysis, or interpretation of data: all authors; drafting of the manuscript: WX, SZ, AY, XZ; critical revision of the

manuscript for important intellectual content: all authors; statistical analysis: WX, SZ, AY, YZ, XZ. All authors contributed to data analysis, drafting and revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Lee ES, Jung SY, Kim JY, et al. Identifying the potential long-term survivors among breast cancer patients with distant metastasis. *Ann Oncol*. 2016;27(5):828–833.
2. Redig AJ, Mcallister SS. Breast cancer as a systemic disease: a view of metastasis. *J Intern Med*. 2013;274(2):113–126.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7–30.
4. van Poznak C, Somerfield MR, Bast RC, et al. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer: american society of clinical oncology clinical practice guideline. *J Clin Oncol*. 2015;33(24):2695–2704.
5. Hoshino A, Costa-Silva B, Shen TL, et al. Tumour exosome integrins determine organotropic metastasis. *Nature*. 2015;527(7578):329–335.
6. Lu X, Kang Y. Organotropism of breast cancer metastasis. *J Mammary Gland Biol Neoplasia*. 2007;12(2-3):153–162.
7. Paget S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev*. 1989;8(2):98–101.
8. Weilbaecher KN, Guise TA, Mccauley LK. Cancer to bone: a fatal attraction. *Nat Rev Cancer*. 2011;11(6):411–425.
9. Kennecke H, Yerushalmi R, Woods R, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol*. 2010;28(20):3271–3277.
10. Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol*. 2010;28(10):1684–1691.
11. Batistatou A, Charalabopoulos A, Charalabopoulos K. Molecular basis of metastasis. *N Engl J Med*. 2009;360(16):1679–1680.
12. Minn AJ, Gupta GP, Siegel PM, et al. Genes that mediate breast cancer metastasis to lung. *Nature*. 2005;436(7050):518–524.
13. Kang Y, Siegel PM, Shu W, et al. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell*. 2003;3(6):537–549.
14. Martin AM, Cagney DN, Catalano PJ, et al. Brain metastases in newly diagnosed breast cancer: a population-based study. *JAMA Oncol*. 2017;3(8):1069–1077.
15. Holm J, Li J, Darabi H, et al. Associations of breast cancer risk prediction tools with tumor characteristics and metastasis. *J Clin Oncol*. 2016;34(3):251–258.
16. Surveillance E, Results E. (SEER) Program Research Data (1973-2013), National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2016, based on the November 2015 submission. Available from: http://www.seer.cancer.gov. Accessed August 15, 2016.
17. Cortesi L, Toss A, Cirilli C, et al. Twenty-years experience with de novo metastatic breast cancer. *Int J Cancer*. 2015;137(6):1417–1426.
18. Kennecke H, Mcarthur H, Olivotto IA, et al. Risk of early recurrence among postmenopausal women with estrogen receptor-positive early breast cancer treated with adjuvant tamoxifen. *Cancer*. 2008;112(7):1437–1444.
19. Kennecke HF, Olivotto IA, Speers C, et al. Late risk of relapse and mortality among postmenopausal women with estrogen responsive early breast cancer after 5 years of tamoxifen. *Ann Oncol*. 2007;18(1):45–51.
20. Dawood S, Broglio K, Esteva FJ, et al. Survival among women with triple receptor-negative breast cancer and brain metastases. *Ann Oncol*. 2009;20(4):621–627.
21. Miller KD, Weathers T, Haney LG, et al. Occult central nervous system involvement in patients with metastatic breast cancer: prevalence, predictive factors and impact on overall survival. *Ann Oncol*. 2003;14(7):1072–1077.
22. Brufsky AM, Mayer M, Rugo HS, et al. Central nervous system metastases in patients with HER2-positive metastatic breast cancer: incidence, treatment, and survival in patients from registHER. *Clin Cancer Res*. 2011;17(14):4834–4843.
23. Clayton AJ, Danson S, Jolly S, et al. Incidence of cerebral metastases in patients treated with trastuzumab for metastatic breast cancer. *Br J Cancer*. 2004;91(4):639–643.
24. Niwińska A, Murawska M, Pogoda K. Breast cancer brain metastases: differences in survival depending on biological subtype, RPA RTOG prognostic class and systemic treatment after whole-brain radiotherapy (WBRT). *Ann Oncol*. 2010;21(5):942–948.
25. Harris E, Barry M, Kell MR. Meta-analysis to determine if surgical resection of the primary tumour in the setting of stage IV breast cancer impacts on survival. *Ann Surg Oncol*. 2013;20(9):2828–2834.
26. Wame S. KPM S, van Gils CH, van der Wall E, Moelans CB, van Diest PJ. Receptor Conversion in Distant Breast Cancer Metastases: A Systematic Review and Meta-analysis. *J Natl Cancer Inst*. 2018;110(6):568–580.