



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/229163/>

Version: Published Version

---

**Article:**

Kong, X., Liu, Q., Qu, Z. et al. (2025) Mapping the metastatic landscape: a population-based cohort study for prognostic insights into newly diagnosed stage IV breast cancer cases. *Cancer Innovation*, 4 (5). e70017. ISSN: 2770-9191

<https://doi.org/10.1002/cai2.70017>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:


<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

ORIGINAL ARTICLE OPEN ACCESS

# Mapping the Metastatic Landscape: A Population-Based Cohort Study for Prognostic Insights Into Newly Diagnosed Stage IV Breast Cancer Cases

Xiangyi Kong<sup>1</sup> | Qiang Liu<sup>1</sup> | Zheng Qu<sup>1</sup>  | Xiangyu Wang<sup>1</sup> | Wenxiang Zhang<sup>1</sup> | Yulu Liu<sup>2</sup> | Robert Coleman<sup>3</sup> | Chunqing Lin<sup>4</sup> | Jing Wang<sup>1</sup>

<sup>1</sup>Department of Breast Surgical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China | <sup>2</sup>Fintech Lab, Department of Computer Science, Chow Yei Ching Building, The University of Hong Kong, Pokfulam, Hong Kong, China | <sup>3</sup>Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK | <sup>4</sup>National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

**Correspondence:** Chunqing Lin ([lincq@cicams.ac.cn](mailto:lincq@cicams.ac.cn)) | Jing Wang ([wangjing@cicams.ac.cn](mailto:wangjing@cicams.ac.cn))

**Received:** 21 July 2024 | **Revised:** 17 February 2025 | **Accepted:** 11 March 2025

**Funding:** This study was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (No. 2023ZD0502602), the CAMS Innovation Fund for Medical Sciences (CIFMS) (No. 2024-I2M-C&T-B-057), the Beijing Natural Science Foundation (Grant No. L234043), and the Natural Science Foundation of China (Grant Nos. 81872160, 82102887, 82103047, and 82173328).

**Keywords:** distant metastasis | distribution and prognosis | newly diagnosed stage IV breast cancer | population-based study | sociodemographic and clinicopathological factor

## ABSTRACT

**Background:** Breast cancer is the most common malignancy and a leading cause of cancer-related deaths among women worldwide. Although treatment advances have improved outcomes, the 5-year survival rate for metastatic breast cancer remains low. Understanding the anatomical distribution, associated risks, and prognostic features of metastases in patients with newly diagnosed stage IV breast cancer is essential for improving clinical management. This study aims to comprehensively investigate these aspects using data from the SEER database.

**Methods:** This study utilized a retrospective cohort design, examining data from the Surveillance, Epidemiology, and End Results (SEER) database. The investigation considered patients diagnosed with stage IV breast cancer from SEER database. Using logistic regression, odds ratios (ORs) were calculated to determine the risk of various metastases, stratified based on sociodemographic and clinicopathological variables. Survival analyses were executed with Kaplan–Meier methodology in tandem with Cox regression analyses.

**Results:** Out of 356,789 breast cancer patients considered, 18,036 (5.06%) were diagnosed with de novo stage IV disease. Bone metastasis predominated with a composition ratio of 42.6%. Patients with the HR–/HER2+ subtype exhibited the highest metastasis incidence at the time of diagnosis, constituting 8.7% of the entire cohort. Male patients displayed heightened susceptibility to bone, lung, and brain metastases compared to female counterparts. Hispanic individuals exhibited the highest propensity for brain metastases. Relative to other subtypes, the HR–/HER2– patients were more inclined toward lung metastases. Those with bone metastasis had a median survival period of 27 months. Grade III patients with brain or liver

**Abbreviations:** AJCC, American Joint Committee on Cancer; CI, confidence interval; FBC, female breast cancer; HER2, human epidermal growth factor receptor 2; HR, hazard ratio HR+/HR–; hormone receptor; IDC, invasive ductal carcinoma; IDLC, invasive ductal and lobular carcinoma; IDM, invasive ductal mammary carcinoma; LC, lobular carcinoma; MBC, male breast cancer; OR, odds ratio; SEER, Surveillance, Epidemiology, and End Results; TNBC, triple-negative breast cancer.

Xiangyi Kong, Qiang Liu, and Zheng Qu contributed equally to this article.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2025 The Author(s). *Cancer Innovation* published by John Wiley & Sons Ltd on behalf of Tsinghua University Press.

metastases faced the most adverse prognoses. A comprehensive profile detailing metastasis patterns by demographics, tumor site and stage, biology, and treatment was presented.

**Conclusions:** This study represents the most comprehensive analysis of metastasis' anatomical distribution and prognosis in breast cancer, offering invaluable insights into metastatic tendencies and characteristics.

## 1 | Introduction

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer-related mortality among women worldwide [1]. The majority of breast cancer patients present with localized disease, primarily confined to the breast and adjacent lymph nodes (stages I–III), with current treatment protocols offering extended survival for most [2, 3]. However, despite significant therapeutic advancements in recent decades, the 5-year survival rate for patients with metastatic breast cancer remains discouragingly low, at approximately 32% [1]. Understanding the characteristics of this critical subgroup is essential for improving clinical surveillance and developing more effective management strategies.

The most common sites of metastasis in breast cancer include the bone, liver, lung, and brain. Several factors, such as tumor size, lymph node involvement, histological grade, pathological subtype, and key biological markers such as hormone receptor and HER2 status, influence the onset and progression of metastatic disease. Wang et al. [4] identified significant differences in distant metastasis patterns among patients with various breast cancer subtypes, noting that those with brain metastasis had the poorest cancer-specific survival across all subtypes. These findings were corroborated by Xiao et al. [5]. Although molecular subtypes play a crucial role in predicting site-specific metastases, recent studies have highlighted the importance of other factors in influencing metastatic patterns. Wang et al. [6] demonstrated that primary tumor location affects metastatic spread, with upper-outer quadrant tumors associating with bone metastases and central tumors showing a higher propensity for liver metastases. Additionally, sociodemographic factors such as race [7] and marital status [8] have been found to modulate metastatic risk. Notably, patients with lower socioeconomic status tend to have higher rates of metastatic disease at diagnosis and are more likely to present with multiple metastatic sites. Despite these insights, the complex interplay between sociodemographic factors, clinicopathological variables, and distant metastasis patterns remains poorly understood. Furthermore, the pattern of multisite metastases, a critical aspect of advanced breast cancer, has yet to be comprehensively explored. This gap underscores the need for more granular investigations into the multifaceted nature of metastatic breast cancer.

In recent years, breast cancer research has expanded beyond the study of localized disease to address the complexities of metastatic breast cancer. This shift reflects an increasing recognition of the urgent need to understand the mechanisms of metastasis, which remains the leading cause of death in breast cancer patients. Leveraging population-based data from the SEER

database, the current study aims (1) to describe the incidence and composite ratio according to major distant metastasis sites, and (2) to investigate the survival and the associations between major clinicopathological variables and different metastasis pattern. By analyzing data from diverse patient demographics and cancer characteristics, we aim to uncover patterns that could inform more personalized and effective treatment approaches.

## 2 | Methods

### 2.1 | Study Design and Data Source

This study employed a population-centric methodology, utilizing data procured from the Surveillance, Epidemiology, and End Results (SEER) database [9]. This approach was conducted in strict adherence to the stipulations delineated by the National Cancer Institute (NCI) under the SEER limited-use data accord.

### 2.2 | Patient Selection

The study focused on primary invasive stage IV breast cancer patients diagnosed within the temporal confines of January 1, 2010, to December 31, 2015. Variables including, but not limited to, gender, insurance codification, and breast tumor laterality, were meticulously extracted. Comprehensive nuances pertaining to data extraction and subsequent analyses are elaborated upon in the Supporting Information.

### 2.3 | Statistical Analysis

A logistic regression framework was used to compute odds ratios (ORs), specifically to evaluate the metastatic risk, stratified based on sociodemographic and clinicopathological determinants. Prognostic evaluations were undertaken utilizing both the Kaplan–Meier and the Cox regression analysis.

In our study, we used logistic regression to analyze the association between various sociodemographic and clinicopathological factors (e.g., age, race, and molecular subtype) and the presence of metastasis at diagnosis. This method was selected because it is suitable for modeling binary outcomes, such as the presence or absence of metastasis (yes/no). The logistic regression outputs provide ORs with associated 95% confidence intervals (CIs) and *p* values, which quantify the strength and direction of these relationships while adjusting for other confounding factors.

For survival analysis, we utilized Cox proportional hazards regression, a widely used model for time-to-event data. This

model was chosen because it accounts for censoring (patients who are lost to follow-up or have not yet experienced the event) and allows for adjustment of multiple covariates, providing hazard ratios (HRs) that represent the relative risk of death associated with different covariates. Both models are commonly used in clinical research to analyze binary outcomes and time-to-event data, and they offer interpretable results that guide clinical decision-making.

To ensure the rigor of our statistical analyses, we provide further details regarding the choice of models and the handling of confounding variables. In our multivariable models, we adjusted for a range of potential confounders, including age, race, molecular subtype, AJCC T and N stage, and treatment modalities. These covariates were selected based on their known associations with metastatic spread and survival outcomes. We also checked for multicollinearity to ensure that no redundant variables were included in the models, thereby avoiding confounding due to highly correlated predictors.

For handling missing data, we applied multiple imputation method to address missing values in key variables.

All statistical results are presented with full transparency. CIs and  $p$  values are provided for all key outcomes, including ORs from logistic regression and HRs from Cox regression. To reduce potential false positive findings due to large sample size, we considered  $p < 0.01$  as statistically significant and offer full explanations of how these  $p$  values, and CIs should be interpreted within the context of our findings. This approach ensures that our statistical methods are transparent, reproducible, and in line with accepted standards for clinical research.

## 2.4 | Ethical Considerations

All operational protocols adhered scrupulously to the ethical guidelines delineated in the 1964 Helsinki declaration [10] and its subsequent amendments [11, 12]. The analytical paradigm was aligned with the stipulations of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [13]. This study has been reported in line with the STROCSS criteria [14].

## 3 | Results

Our comprehensive analysis of 356,789 breast cancer patients provides significant insights into the patterns, risk factors, and prognosis of metastatic breast cancer. This study, one of the largest of its kind, reveals important trends in metastatic site preferences across different molecular subtypes and demographic groups. We uncovered distinct relationships between patient characteristics, tumor biology, and metastatic behavior, which have significant implications for survival outcomes. Our findings highlight the complex interplay between tumor grade, race, age, and metastatic patterns, emphasizing the need for personalized approaches in managing metastatic breast cancer. These results not only confirm some previously known associations but also

reveal novel patterns that could inform future research and clinical decision-making in breast cancer care.

The final cohort incorporated 356,789 breast cancer patients (HR+/HER2– 242,541 [68.0%], HR+/HER2+ 35,090 [9.8%], HR–/HER2+ 15,062 [4.2%], TNBC 37,455 [10.5%]), out of which 18,036 (5.06%) were diagnosed with stage IV of the disease. Notably, the molecular subtype distributions align closely with previous scholarly studies, particularly with those concerning US-based patient cohorts [15]. The data set's attributes are bifurcated and elucidated at two analytical levels: the incidence proportion and the composition ratio, examined both across the entire cohort and specifically within the metastatic subset.

### 3.1 | Incidence Proportion

Of the patients diagnosed with stage IV breast cancer, 13,299 manifested bone metastases (comprising 73.7% of this subgroup), 1519 showcased brain metastases (8.4%), 5182 were diagnosed with liver metastases (28.7%), and 6380 had lung metastases (35.4%). Notably, a fraction of these patients exhibited multisite afflictions. An exhaustive tabulation detailing the numerical and incidence proportions of breast cancer patients with diverse metastatic sites is available in Table S1. These data are further stratified based on socio-demographic and clinicopathological parameters. Principal discoveries from this dataset are comprehensively presented in Table 1A.

### 3.2 | Composition Ratio Analysis

Tables 1 and S2 encapsulate both the count and composition ratios of patients manifesting specific metastatic patterns, categorized by sociodemographic and clinicopathological variables. Table S1 remains agnostic to the differentiation of metastatic site numbers, whereas Table S2 more acutely differentiates metastatic patterns based on the quantity of metastatic sites. Furthermore, the composition ratios of patients within particular sociodemographic and clinicopathological classifications—subdivided by metastatic patterns (with distinct site numbers)—were dissected across the entire breast cancer patient cohort and the subset displaying any form of metastasis. The dataset, however, remains undisclosed. In the metastatic cohort, isolated bone metastasis dominated in prevalence (42.56%), followed by sole lung metastasis (12.11%). The combination of bone and lung metastasis was third (10.87%), whereas the amalgamation of “brain, liver, and lung metastasis” was the least prevalent (0.33%), as detailed in Table S2. The salient observations have been articulated in Table 1B.

### 3.3 | Logistic Regression Analysis of Variables Linked to Specific Metastasis Patterns

Logistic regression examinations were performed across the entire cohort (Table 2) as well as the subset with any form of metastasis (Table S3). The prominent findings are illustrated in Table 1C.

**TABLE 1** | Key findings of metastatic disease in newly diagnosed stage IV breast cancer.

Clinical items	Statistical results and major findings
<b>(A) Incidence proportions for breast cancer metastasis</b>	
This section presents the proportion of patients with each type of metastasis (bone, brain, liver, lung) at the time of diagnosis. These proportions were calculated based on the entire cohort of stage IV breast cancer patients. Statistical analysis was performed using logistic regression to estimate the odds ratios (ORs) for each metastatic site, adjusted for sociodemographic and clinicopathological variables. Results are presented with 95% confidence intervals (CIs) and <i>p</i> values for each category.	
<b>Patient demographics</b>	Among the entire cohort of patients with metastases, patients under 40 had the highest incidence proportion (6.6%), followed by patients over 80 (6.4%), whereas patients aged 40–49 years old had the lowest incidence (4.1%) compared to other age groups. Among the entire cohort, black Americans had the highest incidence of distant metastases (7.5%), followed by white Americans (4.8%) and Hispanics (4%), whereas Asian/Pacific Islanders had the lowest incidence (4.0%) (Table 2). However, among the metastatic cohort, white Americans had the highest incidence (75.4%), and black Americans the lowest (68.4%) for bone metastasis specifically, whereas for brain metastasis, Hispanics had the highest (9.9%) and white Americans had the lowest (7.9) (Table 2). This suggested differences in the distribution of breast cancer metastasis profiles between different races.
<b>Tumor site and stage</b>	Among entire cohort, for all metastasis patterns, the larger the tumor and the higher the AJCC N stage, the higher the incidence of breast cancer metastasis. Patients recorded as having bilateral disease ( <i>n</i> = 1256) appeared to have a very high incidence of metastasis at diagnosis (77.8%), especially in bone (57.2%). Also, among the entire cohort, for all patients with metastases, the incidence for patients with central breast tumor was the highest in terms of primary tumor site (6.05%), followed by tumors that overlapped more than one quadrant (4.11%); the incidence was the lowest for upper-inner quadrant tumor (2.41%).
<b>Tumor biology and treatment</b>	Among the entire cohort, regardless of the site, metastasis incidences were positively correlated with the histology grade. Among the metastatic cohort, brain, liver, and lung metastases incidences were positively correlated with grade, whereas bone metastases showed a negative correlation (Table S1). Among the entire cohort, patients of HR –/HER2+ subtype had the highest incidence proportions for brain metastasis (1.16%), liver metastasis (4.55%), and lung metastasis (3.60%), whereas HR+/HER2+ patients had the highest incidence proportion for bone metastasis (5.16%) (Table S1). The metastasis incidence was lower in patients who had undergone surgery, radiotherapy, or chemotherapy than in patients who had not (all <i>p</i> < 0.01).
<b>(B) Composition ratio analysis for breast cancer metastasis</b>	
This section shows the composition ratios of patients with specific metastatic patterns, stratified by clinical variables. Multivariable logistic regression was used to determine how clinical characteristics, such as age, race, and molecular subtype, impact the likelihood of different metastatic patterns.	
<b>Patient demographics</b>	Among the entire cohort, the proportion of patients aged 60–69 was the largest (26.1%), followed by those aged 50–59 (23.8%), and the third was aged 70–79 (18.0%). The proportion of patients < 40 years old was the least (6.1%) (Table 2). The male–female ratio among the entire cohort was 0.8% versus 99.2%, whereas among the metastatic cohort, it was 1.2% versus 98.8%, which suggested that the male breast cancer (MBC) patients might be more likely to develop metastasis than female breast cancer (FBC) patients (Table 2). Figure S1A showed the ethnic composition ratios in cohorts of different metastatic patterns (white the largest, followed by black, and then Hispanic).
<b>Tumor site and stage</b>	Bilateral breast cancer accounted for 0.4% and 5.4% among the entire and metastatic cohorts, respectively. The patients whose primary breast tumor was located in the upper-outer quadrant accounted for the highest proportion (21.7%) among the entire cohort, overlapping the second highest (18.3%), central the third (6.0%), upper-inner the fourth (5.7%), lower-outer the fifth (4.9%), and lower-inner the sixth (3.2%) (Table 2). Among the metastatic cohort, T4 cancer accounted for the largest proportion (28.7%), followed by T2 (26.3%). The proportion of N1 stage patients accounted for the largest (41.1%), followed by N0 (24.5%), the third was N3 (11.9%), and the fourth was N2 (9.8%) (Table 2 and Figure S1B).

(Continues)

TABLE 1 | (Continued)

Clinical items	Statistical results and major findings
<b>Tumor biology</b>	HR+/HER2- accounted for the largest proportion (51.3%), followed by HR+ /HER2+ (13.9%), third was HR-/HER2- (11.0%), and fourth was HR-/HER2+ (7.3%) in regardless of metastasis pattern (Table 2 and Figure S1C). Grade III had the largest proportion (35.3%), grade II the second (31.4%), grade I the third (5.9%), and grade IV the least (0.7%) (Table 2 and Figure S1D).
<b>(C) Variables associated with specific metastasis pattern by logistic regression</b>	
In this section, we present the odds ratios and statistical significance for various factors (age, sex, and molecular subtype) influencing metastasis at diagnosis. These were derived using logistic regression, adjusting for potential confounders.	
<b>Patient demographics</b>	Among the entire cohort, FBC under 40 were more likely to develop bone metastases and liver metastases and were less likely to develop lung metastases (Table 3). Compared to MBC patients, female patients were less susceptible to bone metastasis (OR 0.71; $p = 0.001$ ) and lung metastasis (OR 0.62; $p < 0.001$ ), but were more likely to develop liver metastasis (OR 2.01; $p < 0.001$ ), whereas no statistical difference for brain metastases was observed ( $p = 0.276$ ) (Table 3). Compared to other races, white patients were the most susceptible to bone metastases (all $p < 0.01$ ). Black patients, however, were more likely to develop lung metastases than white patients (OR 1.15; $p = 0.001$ ; Table 3). No significant differences were found among races for brain metastases. Among the metastatic cohort (Table S3), blacks were the least likely to develop bone metastases (vs. white, OR 0.74; $p < 0.001$ ).
<b>Tumor site and stage</b>	Among the entire cohort (Table 3), bilateral breast cancer was more likely to develop metastases (vs. left, OR 6; $p < 0.001$ ), especially bone metastases (vs. left, OR 2.73; $p < 0.001$ ). Among the metastatic cohort (Table S3), bilateral breast cancer was less likely to develop lung metastasis (vs. left, OR 0.7; $p = 0.001$ ). Among the entire cohort (Table 3), patients with upper-quadrant tumors were less likely to develop distant metastases (especially lung metastases) than those with lower-quadrant tumors ( $p < 0.01$ ). Patients with central tumors were more likely to develop bone metastases (vs. upper-outer, OR 1.31; $p < 0.001$ ). The higher the AJCC T stage of the tumor, the more likely the patients were to develop metastasis ( $p < 0.001$ ; Tables 3 and S3).
<b>Tumor biology</b>	Among the entire cohort, Grade I patients were the least likely to develop metastatic diseases (Table 3). Among the metastatic cohort (Table S3), grade II patients were the most likely to develop bone metastases (vs. grade I, OR 1.1; $p < 0.001$ ); and grade IV patients were the least likely to develop bone metastasis (vs. grade I, OR 0.51; $p = 0.004$ ). As for the pathology, as shown in Table 3, LC was more likely to have bone metastasis (OR 1.19; $p < 0.001$ ) and brain metastasis (OR 1.7; $p < 0.001$ ); but less likely for lung metastasis (OR 0.3; $p < 0.001$ ). LC, IDLC and IDM patients were all less likely to develop lung metastasis than IDC patients (LC: OR 0.3, $p < 0.001$ ; IDLC: OR 0.57, $p < 0.001$ ; and IDM: OR 0.71, $p = 0.009$ ; Table 3). No differences were found for the propensities to develop brain metastases among the several major pathological types. As shown in Table 3, among the entire cohort, for molecular subtype, HR+/HER2- and HR+/HER2+ were more susceptible to developing both bone metastasis and lung metastasis than HR-/HER2+ and HR-/HER2-. HR+/HER2- was the least likely to develop brain metastases and liver metastases. From Table S3 based on the metastatic cohort, the difference of the predisposition to bone metastasis among all the molecular subtypes was not statistically significant. HR+/HER2- subtype was still the least likely to develop brain and liver metastasis.
<b>Treatment</b>	From Tables 3 and S3, patients who had not ever undergone surgery were more likely to develop metastases. Interestingly, compared to patients who had not received radiotherapy, patients who had received radiotherapy were more likely to develop bone metastasis (OR 3.23; $p < 0.001$ ) and brain metastasis (OR 3.75; $p < 0.001$ ; Table 3), but less likely to develop lung metastasis (OR 0.73; $p = 0.007$ ; Table 3). Compared to patients who did not receive chemotherapy, patients who had received chemotherapy were more likely to develop liver metastases (OR 1.57; $p < 0.001$ ; Table 3) and lung metastases (OR 1.16; $p < 0.001$ ; Table 3), but were less likely to develop bone metastases (OR 0.87; $p < 0.001$ ; Table 3).

(Continues)

TABLE 1 | (Continued)

Clinical items	Statistical results and major findings
<b>(D) Survival analysis for breast cancer metastasis</b>	
Kaplan–Meier survival curves and Cox proportional hazards regression were used to evaluate the survival times based on the presence of specific metastatic sites. The section includes hazard ratios (HRs) and <i>p</i> values for each metastatic site and clinical factor.	
<b>Patient demographics</b>	From Cox regression (Table S5), regardless of the metastasis site, generally the older the age, the worse the prognosis. For patients with non-metastatic breast cancer, FBC survived longer than MBC (HR 0.8; <i>p</i> < 0.001), whereas for patients with metastasis, there was no statistical prognosis difference between FBC and MBC, regardless of the metastasis site (Table S5). From Table S5, the black patients had the highest HR to develop metastatic diseases than other races.
<b>Tumor site and stage</b>	The prognosis of right-side breast cancer patients with lung metastasis was better than left-side patients (HR 0.92; <i>p</i> = 0.007; Table S5). In the subset of patients with metastasis at any site, the death risk of patients with bilateral tumors was lower than that of unilateral patients (HR 0.3; <i>p</i> < 0.001), especially for bone metastasis (HR 0.79; <i>p</i> < 0.001; Table S5). Among the metastatic cohort, lower-outer breast cancer patients had a better prognosis than upper-inner ones (HR 0.9; <i>p</i> = 0.045; Table S5). Generally, the higher the T stage, the higher the HR (Table S5). For bone, brain, or lung metastases, the prognosis was not associated with the N stage (Table S5). For liver metastases, N0 stage patients had the highest HR, followed by N2, significantly higher than N1 and N3 patients (Table S5).
<b>Tumor biology and treatment</b>	Among patients with brain metastases or liver metastases, grade III patients had the worst prognosis (brain: grade III vs. grade I, HR 1.69; 95% CI 1.17–2.44; <i>p</i> = 0.005; liver: grade III vs. grade I, HR 1.48; 95% CI 1.21–1.8; <i>p</i> < 0.001); there were no significant differences among grades I, II, and IV patients. Among patients with breast cancer bone metastases, the prognosis of LC was worse than that of IDC (HR 1.11; <i>p</i> = 0.003; Table S5). If the HR was arranged in descending order, in the subset of patients with metastasis at any site, it was HR–/HER2– > HR–/HER2+ > HR+/HER2– > HR+/HER2+ (most <i>p</i> < 0.01; Table S5).
<b>Metastasis pattern</b>	The univariate Cox regression analysis was also conducted based on the four metastasis patterns (data not shown). Overall, breast cancer patients with brain metastasis had the highest HR, suggesting the shortest survival, whereas patients with bone metastasis had the lowest HR, suggesting the longest survival and the best prognosis. The sequence of prognosis of patients with liver and lung metastasis varied according to sociodemographic and clinicopathological variables, such as the molecular subtype.

### 3.4 | Prognostic Analysis

Table S4 furnishes a detailed portrayal of median survival durations for breast cancer patients displaying specific metastatic patterns, further categorized by sociodemographic and clinicopathological variables. Among these, patients with bone metastasis reported the longest median survival duration (27 months), succeeded by lung (18 months), liver (15 months), and brain metastasis (9 months). The Kaplan–Meier evaluations were bifurcated into two distinct paradigms: (1) construction of survival curves for patients, categorized within specific socio-demographic and clinicopathological variables, and delineated by metastatic patterns; (2) depiction of survival curves for patients, characterized by certain metastatic patterns, and further stratified by sociodemographic and clinicopathological criteria. Both Figures 1 and S2 are emblematic of the aforementioned paradigms. Within the subset manifesting any metastasis, a direct correlation was observed: as the count of metastatic sites escalated, the median survival duration plummeted. Typically, increased age corresponded to reduced survival durations, as substantiated by Table S4. Supplementary Kaplan–Meier

visualizations remain undisclosed. The seminal observations are epitomized in Table 1D.

### 3.5 | Subgroup Analysis

Subgroup analysis revealed significant associations between histological grade, race, and metastatic patterns in breast cancer patients. These factors not only influenced the distribution of metastatic sites but also had substantial impacts on survival outcomes. The analysis uncovered complex relationships between tumor biology, patient demographics, and disease progression, highlighting the importance of considering these factors in clinical decision-making and prognostic assessments.

Histological grade analysis (Figures S3 and S4) revealed significant associations with metastatic patterns and survival outcomes in breast cancer patients. As shown in Figure S3, grade I tumors predominantly metastasized to bone (61.5%), whereas higher grade tumors showed increased propensity

**TABLE 2** | Characteristics and composition ratios of breast cancer patients with metastasis [*n* (%)].

<b>Case number (<i>n</i>)</b>	<b>All patients <i>n</i> = 356,789 (%)</b>	<b>No metastasis <i>n</i> = 338,753 (%)</b>	<b>All metastasis <i>n</i> = 18,036 (%)</b>	<b>Bone metastasis <i>n</i> = 13,299 (%)</b>	<b>Brain metastasis <i>n</i> = 1519 (%)</b>	<b>Liver metastasis <i>n</i> = 5182 (%)</b>	<b>Lung metastasis <i>n</i> = 6380 (%)</b>
<b>Diagnosis year</b>							
2010	55763 (15.6)	53012 (15.6)	2751 (15.3)	2001 (15.0)	242 (15.9)	781 (15.1)	937 (14.7)
2011	57981 (16.3)	55098 (16.3)	2883 (16.0)	2148 (16.2)	249 (16.4)	827 (16.0)	1007 (15.8)
2012	59024 (16.5)	56094 (16.6)	2930 (16.2)	2145 (16.1)	216 (14.2)	832 (16.1)	1041 (16.3)
2013	60178 (16.9)	57000 (16.8)	3178 (17.6)	2367 (17.8)	268 (17.6)	953 (18.4)	1124 (17.6)
2014	61110 (17.1)	57954 (17.1)	3156 (17.5)	2324 (17.5)	261 (17.2)	952 (18.4)	1154 (18.1)
2015	62733 (17.6)	59595 (17.6)	3138 (17.4)	2314 (17.4)	283 (18.6)	837 (16.2)	1117 (17.5)
<b>Age group</b>							
< 40	16754 (4.7)	15648 (4.6)	1106 (6.1)	815 (6.1)	88 (5.8)	433 (8.4)	308 (4.8)
40–49	54546 (15.3)	52316 (15.4)	2230 (12.4)	1650 (12.4)	206 (13.6)	781 (15.1)	676 (10.6)
50–59	85021 (23.8)	80735 (23.8)	4286 (23.8)	3251 (24.4)	434 (28.6)	1386 (26.7)	1432 (22.4)
60–69	96800 (27.1)	92091 (27.2)	4709 (26.1)	3551 (26.7)	443 (29.2)	1257 (24.3)	1756 (27.5)
70–79	65422 (18.3)	62171 (18.4)	3251 (18.0)	2368 (17.8)	236 (15.5)	788 (15.2)	1211 (19.0)
≥ 80	38246 (10.7)	35792 (10.6)	2454 (13.6)	1664 (12.5)	112 (7.4)	537 (10.4)	997 (15.6)
<b>Sex</b>							
Male	2799 (0.8)	2588 (0.8)	211 (1.2)	170 (1.3)	20 (1.3)	37 (0.7)	97 (1.5)
Female	353990 (99.2)	336165 (99.2)	17825 (98.8)	13129 (98.7)	1499 (98.7)	5145 (99.3)	6283 (98.5)
<b>Race</b>							
White	245066 (68.7)	233217 (68.8)	11849 (65.7)	8931 (67.2)	933 (61.4)	3324 (64.1)	3957 (62.0)
Black	39113 (11.0)	36166 (10.7)	2947 (16.3)	2017 (15.2)	283 (18.6)	926 (17.9)	1181 (18.5)
Hispanic	39271 (11.0)	37369 (11.0)	1902 (10.5)	1382 (10.4)	188 (12.4)	520 (10.0)	716 (11.2)
Asian/Pacific Islander	29564 (8.3)	28377 (8.4)	1187 (6.6)	863 (6.5)	105 (6.9)	371 (7.2)	469 (7.4)
Other	3775 (1.1)	3624 (1.1)	151 (0.8)	106 (0.8)	10 (0.7)	41 (0.8)	57 (0.9)
<b>Marital status</b>							
None-single	284614 (79.8)	271469 (80.1)	13145 (72.9)	9670 (72.7)	1074 (70.7)	3667 (70.8)	4554 (71.4)
Single	52107 (14.6)	48221 (14.2)	3886 (21.5)	2913 (21.9)	357 (23.5)	1212 (23.4)	1446 (22.7)
Unknown	20068 (5.6)	19063 (5.6)	1005 (5.6)	716 (5.4)	88 (5.8)	303 (5.8)	380 (6.0)
<b>Insurance</b>							
Uninsured	5974 (1.7)	5232 (1.5)	742 (4.1)	554 (4.2)	95 (6.3)	254 (4.9)	332 (5.2)
Insured	343570 (96.3)	326699 (96.4)	16871 (93.5)	12444 (93.6)	1386 (91.2)	4811 (92.8)	5891 (92.3)
Unknown	7245 (2.0)	6822 (2.0)	423 (2.3)	301 (2.3)	38 (2.5)	117 (2.3)	157 (2.5)
<b>Laterality</b>							
Left	180309 (50.5)	171572 (50.6)	8737 (48.4)	6424 (48.3)	715 (47.1)	2519 (48.6)	3104 (48.7)
Right	175083 (49.1)	166839 (49.3)	8244 (45.7)	6102 (45.9)	685 (45.1)	2366 (45.7)	2963 (46.4)
Bilateral	1256 (0.4)	279 (0.1)	977 (5.4)	718 (5.4)	107 (7.0)	279 (5.4)	286 (4.5)
Unknown	141 (0.0)	63 (0.0)	78 (0.4)	55 (0.4)	12 (0.8)	18 (0.3)	27 (0.4)
<b>Primary site</b>							
Upper-outer	117949 (33.1)	114039 (33.7)	3910 (21.7)	2884 (21.7)	293 (19.3)	1099 (21.2)	1237 (19.4)
Upper-inner	42591 (11.9)	41564 (12.3)	1027 (5.7)	751 (5.6)	65 (4.3)	287 (5.5)	343 (5.4)
Lower-inner	19715 (5.5)	19134 (5.6)	581 (3.2)	413 (3.1)	42 (2.8)	147 (2.8)	205 (3.2)

(Continues)

TABLE 2 | (Continued)

<b>Case number (n)</b>	<b>All patients n = 356,789 (%)</b>	<b>No metastasis n = 338,753 (%)</b>	<b>All metastasis n = 18,036 (%)</b>	<b>Bone metastasis n = 13,299 (%)</b>	<b>Brain metastasis n = 1519 (%)</b>	<b>Liver metastasis n = 5182 (%)</b>	<b>Lung metastasis n = 6380 (%)</b>
Lower-outer	26088 (7.3)	25212 (7.4)	876 (4.9)	639 (4.8)	61 (4.0)	239 (4.6)	296 (4.6)
Overlapping	80300 (22.5)	77000 (22.7)	3300 (18.3)	2446 (18.4)	249 (16.4)	960 (18.5)	1175 (18.4)
Central	17939 (5.0)	16854 (5.0)	1085 (6.0)	830 (6.2)	65 (4.3)	278 (5.4)	354 (5.5)
Breast_NOS	48398 (13.6)	41394 (12.2)	7004 (38.8)	5146 (38.7)	714 (47.0)	2109 (40.7)	2682 (42.0)
Other	3809 (1.1)	3556 (1.0)	253 (1.4)	190 (1.4)	30 (2.0)	63 (1.2)	88 (1.4)
<b>Grade</b>							
Grade I	76486 (21.4)	75417 (22.3)	1069 (5.9)	909 (6.8)	49 (3.2)	177 (3.4)	265 (4.2)
Grade II	148828 (41.7)	143160 (42.3)	5668 (31.4)	4578 (34.4)	390 (25.7)	1405 (27.1)	1780 (27.9)
Grade III	108143 (30.3)	101781 (30.0)	6362 (35.3)	4167 (31.3)	585 (38.5)	2178 (42.0)	2565 (40.2)
Grade IV	1184 (0.3)	1061 (0.3)	123 (0.7)	63 (0.5)	16 (1.1)	37 (0.7)	64 (1.0)
Unknown	22148 (6.2)	17334 (5.1)	4814 (26.7)	3582 (26.9)	479 (31.5)	1385 (26.7)	1706 (26.7)
<b>Pathology</b>							
IDC	263582 (73.9)	252160 (74.4)	11422 (63.3)	8221 (61.8)	945 (62.2)	3472 (67.0)	4342 (68.1)
LC	32625 (9.1)	30840 (9.1)	1785 (9.9)	1625 (12.2)	73 (4.8)	323 (6.2)	227 (3.6)
IDLC	19519 (5.5)	18807 (5.6)	712 (3.9)	608 (4.6)	38 (2.5)	157 (3.0)	152 (2.4)
IDM	11572 (3.2)	11341 (3.3)	231 (1.3)	173 (1.3)	18 (1.2)	57 (1.1)	77 (1.2)
Mucinous	6757 (1.9)	6654 (2.0)	103 (0.6)	63 (0.5)	9 (0.6)	21 (0.4)	57 (0.9)
Tubular	1805 (0.5)	1801 (0.5)	4 (0.0)	3 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)
DCM	1402 (0.4)	1357 (0.4)	45 (0.2)	32 (0.2)	3 (0.2)	7 (0.1)	13 (0.2)
Other	19527 (5.5)	15793 (4.7)	3734 (20.7)	2574 (19.4)	433 (28.5)	1144 (22.1)	1512 (23.7)
<b>AJCC T stage</b>							
T1	205136 (57.5)	203224 (60.0)	1912 (10.6)	1430 (10.8)	173 (11.4)	528 (10.2)	529 (8.3)
T2	103812 (29.1)	99074 (29.2)	4738 (26.3)	3522 (26.5)	307 (20.2)	1331 (25.7)	1410 (22.1)
T3	21446 (6.0)	18902 (5.6)	2544 (14.1)	1831 (13.8)	182 (12.0)	700 (13.5)	871 (13.7)
T4	14401 (4.0)	9220 (2.7)	5181 (28.7)	3776 (28.4)	480 (31.6)	1532 (29.6)	2375 (37.2)
TX	10520 (2.9)	7279 (2.1)	3241 (18.0)	2429 (18.3)	332 (21.9)	996 (19.2)	1070 (16.8)
Tis	492 (0.1)	492 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
T0	959 (0.3)	562 (0.2)	397 (2.2)	304 (2.3)	43 (2.8)	90 (1.7)	107 (1.7)
Unknown	23 (0.0)	0 (0.0)	23 (0.1)	7 (0.1)	2 (0.1)	5 (0.1)	18 (0.3)
<b>AJCC N stage</b>							
N0	242120 (67.9)	237699 (70.2)	4421 (24.5)	3242 (24.4)	378 (24.9)	1195 (23.1)	1422 (22.3)
N1	80727 (22.6)	73316 (21.6)	7411 (41.1)	5508 (41.4)	593 (39.0)	2231 (43.1)	2755 (43.2)
N2	17878 (5.0)	16106 (4.8)	1772 (9.8)	1295 (9.7)	136 (9.0)	479 (9.2)	629 (9.9)
N3	11011 (3.1)	8871 (2.6)	2140 (11.9)	1576 (11.9)	185 (12.2)	583 (11.3)	756 (11.8)
NX	5053 (1.4)	2761 (0.8)	2292 (12.7)	1678 (12.6)	227 (14.9)	694 (13.4)	818 (12.8)
<b>AJCC stage</b>							
I	177092 (49.6)	177091 (52.3)	1 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.0)
II	114720 (32.2)	114720 (33.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
III	38462 (10.8)	38436 (11.3)	26 (0.1)	9 (0.1)	2 (0.1)	6 (0.1)	18 (0.3)
IV	18008 (5.0)	0 (0.0)	18008 (99.8)	13289 (99.9)	1517 (99.9)	5175 (99.9)	6361 (99.7)
0	492 (0.1)	492 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

(Continues)

TABLE 2 | (Continued)

Case number (n)	All patients n = 356,789 (%)	No metastasis n = 338,753 (%)	All metastasis n = 18,036 (%)	Bone metastasis n = 13,299 (%)	Brain metastasis n = 1519 (%)	Liver metastasis n = 5182 (%)	Lung metastasis n = 6380 (%)
Unknown	8015 (2.2)	8014 (2.4)	1 (0.0)	1 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)
<b>Extra metastasis</b>							
No extra sites	338753 (94.9)	338753 (100.0)	0 (0.0)	7472 (56.2)	288 (19.0)	1422 (27.4)	2115 (33.2)
1	11635 (3.3)	0 (0.0)	11635 (64.5)	3972 (29.9)	556 (36.6)	2229 (43.0)	2627 (41.2)
2	4692 (1.3)	0 (0.0)	4692 (26.0)	1416 (10.6)	430 (28.3)	1244 (24.0)	1335 (20.9)
3	1475 (0.4)	0 (0.0)	1475 (8.2)	234 (1.8)	234 (15.4)	234 (4.5)	234 (3.7)
4	234 (0.1)	0 (0.0)	234 (1.3)	205 (1.5)	11 (0.7)	53 (1.0)	69 (1.1)
<b>Subtype</b>							
HR+/HER2–	242541 (68.0)	233288 (68.9)	9253 (51.3)	7642 (57.5)	563 (37.1)	1957 (37.8)	2864 (44.9)
HR+/HER2+	35090 (9.8)	32576 (9.6)	2514 (13.9)	1811 (13.6)	222 (14.6)	992 (19.1)	834 (13.1)
HR–/HER2+	15062 (4.2)	13752 (4.1)	1310 (7.3)	704 (5.3)	175 (11.5)	685 (13.2)	542 (8.5)
HR–/HER2–	37455 (10.5)	35476 (10.5)	1979 (11.0)	1055 (7.9)	274 (18.0)	668 (12.9)	973 (15.3)
Unknown	26641 (7.5)	23661 (7.0)	2980 (16.5)	2087 (15.7)	285 (18.8)	880 (17.0)	1167 (18.3)
<b>Surgery</b>							
Not performed	29462 (8.3)	16444 (4.9)	13018 (72.2)	9852 (74.1)	1290 (84.9)	4008 (77.3)	4866 (76.3)
Performed	324372 (90.9)	319633 (94.4)	4739 (26.3)	3253 (24.5)	217 (14.3)	1103 (21.3)	1424 (22.3)
Unknown	2955 (0.8)	2676 (0.8)	279 (1.5)	194 (1.5)	12 (0.8)	71 (1.4)	90 (1.4)
<b>Radiotherapy</b>							
Not performed	6103 (1.7)	5847 (1.7)	256 (1.4)	177 (1.3)	21 (1.4)	68 (1.3)	119 (1.9)
Performed	170666 (47.8)	165199 (48.8)	5467 (30.3)	4563 (34.3)	919 (60.5)	1183 (22.8)	1577 (24.7)
None/Unknown	180020 (50.5)	167707 (49.5)	12313 (68.3)	8559 (64.4)	579 (38.1)	3931 (75.9)	4684 (73.4)
<b>Chemotherapy</b>							
No/Unknown	217968 (61.1)	208870 (61.7)	9098 (50.4)	6972 (52.4)	728 (47.9)	2108 (40.7)	3179 (49.8)
Yes	138821 (38.9)	129883 (38.3)	8938 (49.6)	6327 (47.6)	791 (52.1)	3074 (59.3)	3201 (50.2)

Note: This table outlines the clinical characteristics and composition ratios of patients with different metastatic sites (bone, brain, liver, and lung) at diagnosis. Variables such as age group, sex, race, and marital status are presented to show how these factors influence the likelihood of metastasis. Statistical analyses were conducted using logistic regression to adjust for confounding factors. The percentages represent the distribution of metastatic cases across different subgroups within the cohort. Abbreviations: DCM, ductal carcinoma, micropapillary; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; IDC, infiltrating duct carcinoma; IDLC, infiltrating duct and lobular carcinoma; IDM, infiltrating duct mixed with other types of carcinoma; LC, lobular carcinoma; Mucinous, mucinous adenocarcinoma; NOS, not otherwise specified; Tubular, tubular adenocarcinoma; +, positive; –, negative.

for lung and liver metastases. Grade IV tumors exhibited the highest rate of lung metastases (26.0%). Interestingly, the proportion of multiorgan metastases ( $\geq 3$  organs) was the highest in grade III tumors instead of grade IV, suggesting a complex relationship between tumor grade and metastatic spread. The Kaplan–Meier survival analysis (Figure S4) demonstrated a clear stratification based on grade, with grade I patients showing the best prognosis and grade IV the worst (below 10% at 60 months). Notably, grades III and IV patients experienced rapid initial declines in survival, indicating a higher risk of early mortality post-diagnosis.

Racial disparities were evident in both metastatic patterns (Figure S5) and survival outcomes (Figure S6). Figure S5 demonstrates that bone was the most common metastatic site across all races, with white patients showing the highest

prevalence (44.9%) and black patients the lowest (36.7%). Lung metastases were most frequent in black patients (15.0%), whereas liver metastases were most common among Asian patients (8.7%). Brain metastases showed less variation among races.

As illustrated in Figure S6, Kaplan–Meier analysis revealed significant differences in overall survival across racial groups ( $p < 0.0001$ ). Asian and Pacific Islander patients demonstrated superior survival rates, whereas black patients faced the most challenging prognosis, with survival rates approximating 15% at 80 months. Notably, the “Other” category displayed the most favorable survival curve; however, this finding warrants cautious interpretation due to the limited sample size. The observed survival disparities persisted across various metastatic patterns, suggesting that race may serve as an independent

**TABLE 3** | Multivariable logistic regression for the presence of metastases at diagnosis of breast cancer (among the entire cohort).

Variable	All metastasis		Bone metastasis		Brain metastasis		Liver metastasis		Lung metastasis	
	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value
<b>Age group</b>										
< 40	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
40–49	0.81 (0.74–0.89)	< 0.001	0.75 (0.67–0.84)	< 0.001	1.13 (0.86–1.49)	0.394	0.79 (0.69–0.92)	0.001	1.14 (0.97–1.34)	0.113
50–59	0.97 (0.89–1.06)	0.461	0.86 (0.78–0.96)	0.005	1.22 (0.95–1.58)	0.134	0.73 (0.64–0.84)	< 0.001	1.42 (1.23–1.65)	< 0.001
60–69	1.01 (0.93–1.11)	0.771	0.83 (0.75–0.92)	< 0.001	1.23 (0.95–1.60)	0.119	0.61 (0.53–0.70)	< 0.001	1.89 (1.63–2.19)	< 0.001
70–79	0.95 (0.87–1.05)	0.324	0.74 (0.66–0.83)	< 0.001	0.96 (0.72–1.27)	0.758	0.58 (0.50–0.67)	< 0.001	2.05 (1.76–2.40)	< 0.001
≥ 80	0.5 (0.45–0.55)	< 0.001	0.44 (0.39–0.49)	< 0.001	0.58 (0.42–0.81)	0.001	0.48 (0.41–0.57)	< 0.001	2.1 (1.79–2.48)	< 0.001
<b>Sex</b>										
Male	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Female	0.83 (0.69–1.01)	0.058	0.71 (0.57–0.88)	0.001	0.76 (0.47–1.28)	0.276	2.01 (1.38–3.01)	< 0.001	0.62 (0.48–0.80)	< 0.001
<b>Race</b>										
White	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Black	0.92 (0.87–0.97)	0.005	0.8 (0.75–0.86)	< 0.001	1.09 (0.93–1.27)	0.275	0.99 (0.90–1.08)	0.782	1.15 (1.06–1.25)	0.001
Hispanic	0.68 (0.63–0.72)	< 0.001	0.71 (0.66–0.77)	< 0.001	1.17 (0.98–1.39)	0.088	0.67 (0.60–0.74)	< 0.001	1.07 (0.97–1.18)	0.164
Asian/Pacific Islander	0.68 (0.62–0.73)	< 0.001	0.68 (0.62–0.75)	< 0.001	0.96 (0.76–1.19)	0.708	0.79 (0.70–0.90)	< 0.001	1.02 (0.91–1.15)	0.689
Other	0.49 (0.39–0.60)	< 0.001	0.54 (0.42–0.69)	< 0.001	0.61 (0.28–1.17)	0.176	0.67 (0.46–0.94)	0.025	0.88 (0.64–1.19)	0.425
<b>Laterality</b>										
Left	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Right	0.97 (0.93–1.01)	0.136	0.98 (0.94–1.03)	0.423	1.02 (0.91–1.15)	0.727	0.97 (0.91–1.03)	0.325	1.01 (0.95–1.07)	0.857
Bilateral	6 (5.02–7.20)	< 0.001	2.73 (2.32–3.22)	< 0.001	1.12 (0.85–1.47)	0.416	1.2 (0.99–1.44)	0.060	0.87 (0.73–1.04)	0.136
Unknown	4.17 (2.56–6.91)	< 0.001	2.08 (1.29–3.37)	0.003	1.89 (0.92–3.62)	0.067	0.88 (0.49–1.52)	0.662	1.19 (0.70–1.97)	0.510
<b>Primary site</b>										
Upper-outer	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Upper-inner	1.02 (0.94–1.11)	0.631	0.98 (0.89–1.08)	0.636	0.83 (0.62–1.1)	0.201	1 (0.86–1.15)	0.961	1.06 (0.92–1.21)	0.432
Lower-inner	1.21 (1.08–1.34)	0.001	1.11 (0.98–1.26)	0.105	1.03 (0.72–1.43)	0.886	0.97 (0.80–1.17)	0.747	1.28 (1.08–1.51)	0.004
Lower-outer	1.16 (1.06–1.27)	0.002	1.11 (1.00–1.23)	0.059	0.95 (0.70–1.27)	0.735	0.99 (0.84–1.15)	0.868	1.22 (1.06–1.41)	0.007
Overlapping	1.1 (1.04–1.16)	0.002	1.07 (1.00–1.15)	0.039	1.06 (0.88–1.27)	0.548	1.05 (0.95–1.16)	0.323	1.15 (1.05–1.26)	0.004
Central	1.32 (1.21–1.45)	< 0.001	1.31 (1.18–1.45)	< 0.001	1.01 (0.75–1.34)	0.969	1.13 (0.97–1.32)	0.112	1.13 (0.98–1.30)	0.097

(Continues)

TABLE 3 | (Continued)

Variable	All metastasis		Bone metastasis		Brain metastasis		Liver metastasis		Lung metastasis	
	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Breast_NOS	1.63 (1.54–1.73)	< 0.001	1.35 (1.26–1.44)	< 0.001	1.38 (1.17–1.63)	< 0.001	1.21 (1.10–1.33)	< 0.001	1.4 (1.28–1.52)	< 0.001
Other	1.34 (1.13–1.59)	0.001	1.24 (1.02–1.51)	0.029	1.97 (1.27–2.96)	0.002	1.06 (0.78–1.42)	0.700	1.29 (0.99–1.66)	0.052
<b>Histological grade</b>										
Grade I	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Grade II	1.5 (1.39–1.62)	< 0.001	1.4 (1.29–1.53)	< 0.001	1.73 (1.27–2.39)	0.001	1.59 (1.35–1.90)	< 0.001	1.48 (1.28–1.71)	< 0.001
Grade III	1.65 (1.52–1.79)	< 0.001	1.3 (1.19–1.43)	< 0.001	2.07 (1.52–2.87)	< 0.001	2.07 (1.75–2.47)	< 0.001	1.83 (1.58–2.12)	< 0.001
Grade IV	2.31 (1.75–3.02)	< 0.001	1.11 (0.77–1.57)	0.585	3.29 (1.65–6.25)	< 0.001	2.03 (1.29–3.11)	0.002	3.08 (2.12–4.41)	< 0.001
Unknown	1.99 (1.81–2.18)	< 0.001	1.7 (1.53–1.88)	< 0.001	1.87 (1.36–2.63)	< 0.001	1.8 (1.50–2.17)	< 0.001	1.62 (1.38–1.90)	< 0.001
<b>Pathology</b>										
IDC	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
LC	1.19 (1.11–1.27)	< 0.001	1.7 (1.59–1.83)	< 0.001	0.79 (0.60–1.01)	0.072	0.84 (0.74–0.96)	0.012	0.3 (0.26–0.35)	< 0.001
IDLC	1.06 (0.96–1.16)	0.254	1.33 (1.20–1.47)	< 0.001	0.86 (0.60–1.19)	0.377	0.89 (0.74–1.06)	0.206	0.57 (0.47–0.68)	< 0.001
IDM	0.69 (0.59–0.81)	< 0.001	0.79 (0.66–0.94)	0.01	0.98 (0.58–1.57)	0.951	0.79 (0.59–1.04)	0.104	0.71 (0.55–0.91)	0.009
Mucinous	0.47 (0.37–0.59)	< 0.001	0.36 (0.26–0.48)	< 0.001	1.16 (0.54–2.2)	0.674	0.66 (0.40–1.02)	0.079	1.19 (0.88–1.58)	0.248
Tubular	0.29 (0.09–0.71)	0.018	0.31 (0.08–0.82)	0.046	0 (0.00–0.00)	0.939	0.49 (0.03–2.27)	0.483	0 (0.00–0.00)	0.897
DCM	0.54 (0.37–0.77)	0.001	0.62 (0.40–0.93)	0.026	0.83 (0.20–2.34)	0.766	0.43 (0.18–0.88)	0.037	0.64 (0.33–1.13)	0.150
Other	1.29 (1.20–1.38)	< 0.001	1.02 (0.94–1.11)	0.632	1.29 (1.1–1.52)	0.002	1.1 (0.99–1.22)	0.069	1.09 (0.99–1.20)	0.073
<b>AJCC T stage</b>										
T1	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
T2	2.42 (2.28–2.58)	< 0.001	2.54 (2.36–2.73)	< 0.001	1.27 (1.03–1.56)	0.024	1.83 (1.63–2.04)	< 0.001	2.24 (2.01–2.5)	< 0.001
T3	3.95 (3.67–4.25)	< 0.001	3.47 (3.18–3.78)	< 0.001	1.47 (1.15–1.87)	0.002	2.15 (1.88–2.46)	< 0.001	3.82 (3.37–4.33)	< 0.001
T4	7.87 (7.32–8.47)	< 0.001	5.43 (4.98–5.92)	< 0.001	1.58 (1.27–1.97)	< 0.001	2.21 (1.95–2.51)	< 0.001	5.99 (5.33–6.74)	< 0.001
TX	2.82 (2.60–3.07)	< 0.001	2.95 (2.68–3.25)	< 0.001	1.63 (1.28–2.07)	< 0.001	2.13 (1.85–2.45)	< 0.001	2.71 (2.37–3.10)	< 0.001
<b>AJCC N stage</b>										
N0	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
N1	2.67 (2.54–2.81)	< 0.001	2.29 (2.17–2.43)	< 0.001	1.19 (1.02–1.38)	0.026	1.59 (1.46–1.73)	< 0.001	1.9 (1.75–2.05)	< 0.001
N2	3.77 (3.49–4.07)	< 0.001	3.19 (2.92–3.48)	< 0.001	1.44 (1.15–1.81)	0.002	1.89 (1.66–2.15)	< 0.001	2.19 (1.94–2.46)	< 0.001
N3	6.33 (5.85–6.85)	< 0.001	4.79 (4.38–5.24)	< 0.001	1.64 (1.33–2.02)	< 0.001	1.92 (1.69–2.17)	< 0.001	2.49 (2.22–2.8)	< 0.001
NX	3.56 (3.26–3.88)	< 0.001	2.51 (2.28–2.77)	< 0.001	1.51 (1.23–1.84)	< 0.001	1.87 (1.65–2.12)	< 0.001	2.07 (1.84–2.33)	< 0.001

(Continues)

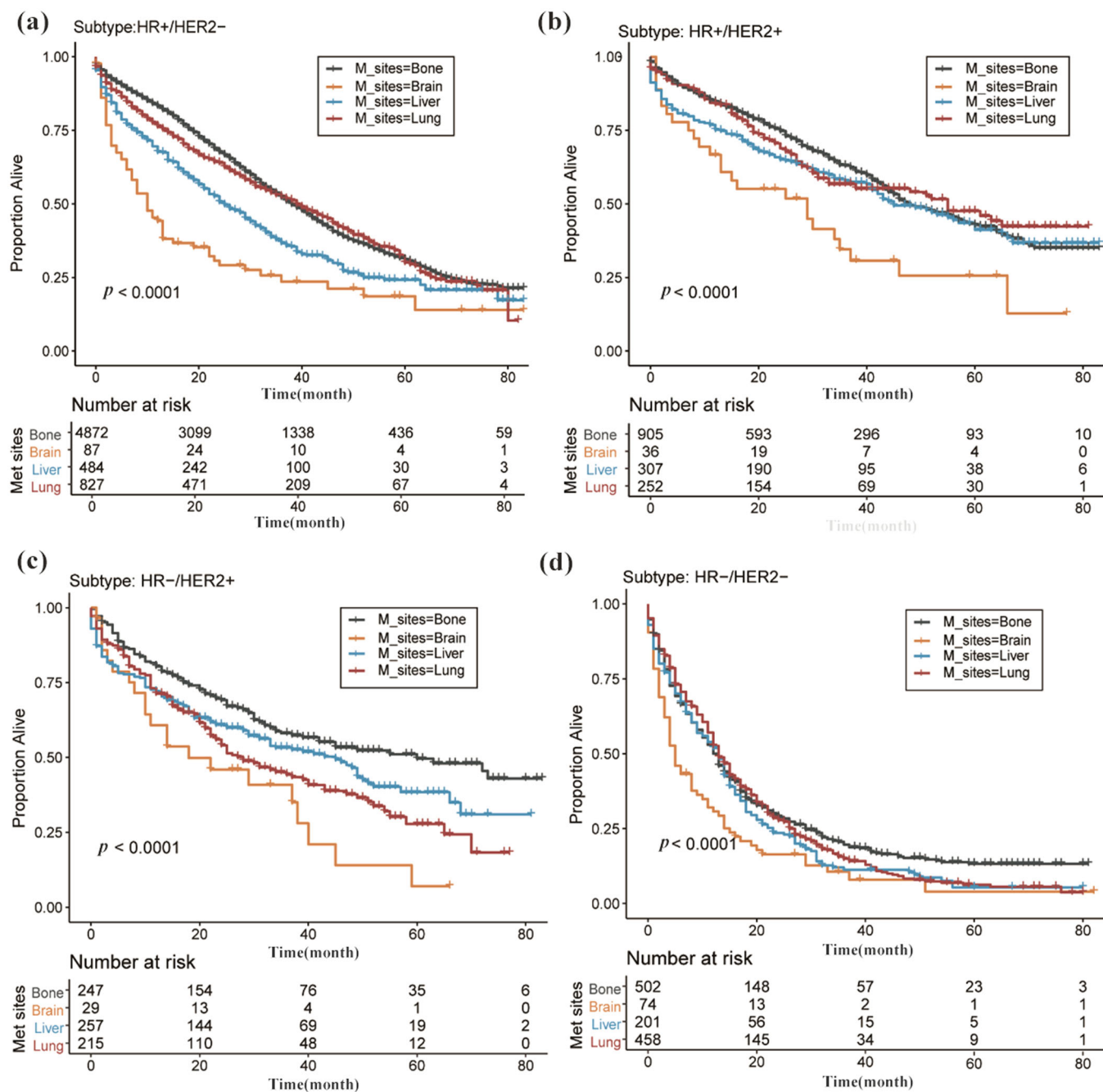
TABLE 3 | (Continued)

Variable	All metastasis		Bone metastasis		Brain metastasis		Liver metastasis		Lung metastasis	
	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value
<b>Metastasis site</b>										
1 site	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
2 sites	NA	NA	5.55 (5.2–5.92)	< 0.001	6.69 (5.51–8.15)	< 0.001	8.45 (7.73–9.24)	< 0.001	6.36 (5.90–6.87)	< 0.001
3 sites	NA	NA	9.03 (8.04–10.15)	< 0.001	11.2 (9.12–13.8)	< 0.001	16.83 (15.1–18.75)	< 0.001	12.48 (11.29–13.78)	< 0.001
4 sites	NA	NA	14.76 (10.6–20.89)	< 0.001	20.13 (15.87–25.57)	< 0.001	37.29 (29.93–46.48)	< 0.001	37.41 (29.24–48.07)	< 0.001
<b>Molecular subtype</b>										
HR+/HER2–	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
HR+/HER2+	1.22 (1.15–1.30)	< 0.001	0.94 (0.87–1.01)	0.086	1.6 (1.34–1.91)	< 0.001	2.12 (1.93–2.33)	< 0.001	1 (0.91–1.10)	0.995
HR–/HER2+	1.1 (1.01–1.19)	0.030	0.45 (0.40–0.51)	< 0.001	2.64 (2.15–3.24)	< 0.001	3.21 (2.86–3.60)	< 0.001	1.23 (1.09–1.39)	0.001
HR–/HER2–	0.81 (0.75–0.87)	< 0.001	0.42 (0.38–0.46)	< 0.001	2.58 (2.16–3.07)	< 0.001	1.41 (1.26–1.57)	< 0.001	1.49 (1.35–1.64)	< 0.001
Unknown	0.82 (0.77–0.88)	< 0.001	0.68 (0.62–0.73)	< 0.001	1.57 (1.31–1.87)	< 0.001	1.38 (1.24–1.53)	< 0.001	1.13 (1.03–1.25)	0.013
<b>Surgery</b>										
Not performed	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Performed	0.03 (0.03–0.03)	< 0.001	0.04 (0.04–0.04)	< 0.001	0.06 (0.05–0.07)	< 0.001	0.14 (0.13–0.15)	< 0.001	0.15 (0.14–0.16)	< 0.001
Unknown	0.11 (0.09–0.12)	< 0.001	0.2 (0.17–0.24)	< 0.001	0.3 (0.16–0.52)	< 0.001	0.3 (0.23–0.38)	< 0.001	0.34 (0.27–0.42)	< 0.001
<b>Radiotherapy</b>										
Not performed	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Radiotherapy performed	2.76 (2.33–3.27)	< 0.001	3.23 (2.66–3.95)	< 0.001	3.75 (2.41–6.14)	< 0.001	0.76 (0.58–1.02)	0.059	0.73 (0.58–0.92)	0.007
None/unknown	2.28 (1.94–2.69)	< 0.001	1.93 (1.59–2.35)	< 0.001	0.7 (0.45–1.14)	0.125	1.55 (1.19–2.06)	0.002	1.19 (0.95–1.49)	0.131
<b>Chemotherapy</b>										
No/Unknown	[Reference]		[Reference]		[Reference]		[Reference]		[Reference]	
Yes	1.01 (0.96–1.06)	0.656	0.87 (0.82–0.92)	< 0.001	0.88 (0.77–1.00)	0.044	1.57 (1.46–1.70)	< 0.001	1.16 (1.08–1.24)	< 0.001

*Note:* This table presents the results of multivariable logistic regression analysis examining the relationship between various clinicopathological factors (such as age, race, molecular subtype, and treatment modalities) and the presence of metastasis at diagnosis. For each metastatic site (bone, brain, liver, and lung), the odds ratios (ORs) and their 95% confidence intervals (CIs) are reported, along with *p* values for statistical significance. These results were adjusted for potential confounding variables to ensure the accuracy and reliability of the estimates. OR (95% CI): The odds ratio (OR) reflects the likelihood of a given factor being associated with metastasis at diagnosis. A 95% CI is provided to show the precision of the OR estimate. The *p* value indicates the statistical significance of each variable in relation to metastasis presence. A *p* value of < 0.01 is considered statistically significant.

Marital status and insurance status were also included in multivariate analyses but not shown due to significance.

Abbreviations: CI, confidence interval; DCM, ductal carcinoma, micropapillary; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; IDC, infiltrating duct carcinoma; IDLC, infiltrating duct and lobular carcinoma; IDM, infiltrating duct mixed with other types of carcinomas; LC, lobular carcinoma; Mucinous, mucinous adenocarcinoma; NA, not applicable; NOS, not otherwise specified; OR, odds ratio; Tubular, tubular adenocarcinoma; +, positive; –, negative.



**FIGURE 1** | Kaplan–Meier survival curve for patients stratified by the metastasis pattern. Survival curves of four different metastasis patterns for breast cancer patients of HR+/HER2– (a), HR+/HER2+ (b), HR–/HER2+ (c) as well as HR–/HER2– (d). From (a), for HR+/HER2– breast cancer patients, within about 30 months of follow-up, in the order of best to poorest in prognosis, bone metastasis ranked the first, lung metastasis the second, liver metastasis the third, and brain metastasis the last. From the 30th month to the 60th month, the survival curves of bone metastasis and lung metastasis tended to coincide. After the 60th month, the survival curve of liver metastases gradually flattened, approaching bone metastases and lung metastases. After the 80th month, the arrangement of the survival curve order changed: bone metastasis ranked the best, liver metastasis the second, brain metastasis the third, and lung metastasis the last. However, for patients with HR–/HER2+ breast cancer, the distribution of the four curves in the Kaplan–Meier plot was relatively more distinct than the HR+/HER2– subtype: the prognosis of bone metastasis was the best, liver metastasis the second, lung metastasis the third, and brain metastasis the last in all follow-up time (b).

prognostic factor in metastatic breast cancer. These findings underscore the complex interplay between racial factors, metastatic tendencies, and survival outcomes, emphasizing the necessity for personalized therapeutic strategies and prognostic assessments that account for racial distinctions in managing metastatic breast cancer.

#### 4 | Discussion

Prior literature on this topic can generally be categorized into two main approaches: the first focuses on analyzing metastasis patterns and prognosis by isolating individual variables such as subtype or gender. The second emphasizes the role of molecular

subtypes in influencing metastasis distribution and prognosis, often concentrating on a specific metastatic site in breast cancer. Using the SEER dataset from 2010 to 2013, Wang et al. [4] found that bone metastasis was the most prevalent, whereas brain metastasis occurred less frequently. Their study highlighted that the HR+/HER2- subtype was most strongly associated with metastasis, whereas the HR-/HER2+ subtype exhibited the lowest metastasis rates. However, their analysis was limited to the molecular subtypes. Similarly, Xiao et al. [5] explored metastatic patterns and prognoses through the lens of subtypes, identifying a strong correlation between HR+/HER2+ and HR-/HER2+ subtypes and an increased risk of metastasis to the liver, brain, and lungs, whereas the HR-/HER2- subtype was inversely associated with bone metastasis. Furthermore, Xie et al. [16] examined metastatic patterns in male breast cancer (MBC) and compared them with female breast cancer (FBC), revealing distinct clinicopathological features and metastatic patterns in metastatic MBC. However, prognostic outcomes between MBC and FBC patients were similar, despite a higher incidence of central breast tumors and older age ( $\geq 60$  years) in MBC patients. Martin et al. [17] focused on the incidence and prognosis of brain metastases, noting higher incidence rates among patients with HR-/HER2+ and HR-/HER2- subtypes. Interestingly, patients with HR+/HER2+ breast cancer had better survival rates, whereas those with HR-/HER2- had the poorest prognosis. A common limitation across these studies is their narrow focus on a single metastatic site or a single variable, which may overlook essential insights that a more comprehensive, comparative analysis could reveal. Moreover, none of these studies accounted for the role of metastasis relapse.

In contrast to most existing literature, our study takes a broader approach by examining all patterns of distant metastasis in breast cancer patients at the time of diagnosis, utilizing a multifaceted analytical perspective supported by a large sample size. Xiao et al.'s research emphasizes that the impact of age on metastatic patterns is not uniform across different sites [5]. In our study, we observed that the impact of age on metastasis varied significantly depending on the metastatic site. Younger patients ( $< 40$  years) exhibited a higher incidence of bone metastasis, which is consistent with previous studies suggesting that younger women often have higher bone mineral density and a more active bone microenvironment, providing a favorable site for cancer cell growth. Additionally, younger patients may have higher estrogen levels, which can promote bone metastasis in hormone receptor-positive tumors. In contrast, older patients ( $> 70$  years) showed a higher propensity for lung metastasis. This could be due to age-related declines in immune function and the increased vulnerability of the lungs to metastatic disease. Furthermore, as the immune system becomes less efficient with aging, the ability to suppress micrometastases may be reduced, allowing tumors to establish in distant organs such as the lungs and liver. Interestingly, the distribution of metastasis to the brain did not show a significant age-related difference, suggesting that brain metastases may be more closely associated with the tumor's biological characteristics (e.g., HER2+ or triple-negative) rather than age. The data indicate that grade III tumors, which are more common in younger patients, have a higher tendency to metastasize to the brain, whereas liver and lung metastases are often more

prevalent in older patients, reflecting the increased systemic spread and immune system inefficiency in this population.

Notably, younger breast cancer patients seemed predisposed to a heightened risk of bone metastases, although the variance failed to achieve statistical significance. As age advanced, the proclivity for liver metastases diminished, whereas the susceptibility to lung metastases ascended [5]. Pertaining to brain metastases, individuals within the age bracket of 40–65 years manifested a more pronounced risk in comparison to other age cohorts [5]. This alignment of findings mirrors our own: the influence of age exhibited fluctuations depending on the metastasis site. Elevated age cohorts demonstrated an augmented predisposition for lung metastasis, but a reduced risk for bone and liver metastasis. Age-related associations with brain metastasis did not achieve statistical significance. Xie et al. discerned that, in juxtaposition with non-metastatic MBC patients, those with metastatic MBC portrayed distinct clinicopathological characteristics and deviated from their metastatic FBC counterparts. Nonetheless, prognostic outcomes bore no significant distinctions between metastatic MBC and FBC patients. In our exploration, non-metastatic breast cancer female patients exhibited a more extended survival span compared to their male counterparts (HR 0.8; 95% CI 0.73–0.87;  $p < 0.001$ ). Concerning metastatic patients, our analysis discerned that the subtype distribution was roughly analogous between genders. However, male patients presented a diminished composition ratio of TNBC (1.8%) in contrast to female patients (10.6%), a finding in concordance with previously cited data [18].

Interestingly, our study revealed that single (unmarried) patients exhibited an elevated incidence of metastasis. This observation aligns with prior literature indicating that unmarried individuals tend to experience poorer survival outcomes in breast cancer [8]. A potential rationale for this phenomenon might be the heightened susceptibility of unmarried patients to psychological distress and engagement in detrimental habits, stemming from the absence of financial stability and emotional support often provided by spouses [19, 20]. Remarkably, after neutralizing the stage effect in multivariate analyses, the heightened metastasis risk resurfaced. However, when mitigating the effect of subtype, the outcome remained largely unchanged. This indicates that delayed diagnosis could be a pivotal determinant of an augmented metastasis risk among black patients. Furthermore, within the entire sample, the incidence of bone metastasis for bilateral breast cancer surpassed that of unilateral cases (57.17% bilateral vs. 3.56% left). However, among those with metastasis, the lung metastasis incidence for bilateral cancer was relatively diminished in comparison to unilateral cases (29.27% bilateral vs. 35.53% left). Distinctively, patients with bilateral breast cancer manifested inferior survival outcomes compared to their counterparts. Such an observation might be elucidated by the fact that (1) bilateral breast cancer patients tend to receive a diagnosis at more advanced stages than those with unilateral manifestations; and (2) a significant proportion of apparent bilateral primary tumors might actually represent metastatic growths from the opposing side. The reduced metastasis incidence in patients undergoing surgery, radiotherapy, or chemotherapy is presumably inversely correlated, as those manifesting overt metastasis may forgo locoregional treatments. Pertaining to bone metastasis, black

patients, in contrast to their white counterparts, consistently demonstrated a significantly diminished OR for manifesting bone metastasis at diagnosis. There was an absence of disparity between white and black patients concerning brain and lung metastasis, a finding incongruent with the conclusions drawn by Martin et al. [17]. This deviation necessitates further investigations for elucidation. Our data indicated a predisposition for patients undergoing chemotherapy to develop liver metastases, whereas those receiving radiotherapy exhibited a proclivity for bone metastasis. A plausible explanation for such an occurrence might be the common clinical practice of prescribing radiotherapy for bone metastasis and chemotherapy for liver metastasis.

We observed that uninsured patients exhibited a heightened proportion of metastatic incidence compared to those insured. This trend may be attributed to the uninsured cohort's probable association with an absence of preventative measures, limited screening, restricted access to healthcare, and diagnostic delays, culminating in more advanced disease presentations at diagnosis [21]. There is a dearth of studies examining metastatic incidence amongst patients with bilateral breast cancer. Our findings underscored an elevated metastatic incidence amongst this demographic. A singular precedent study highlighted that patients with bilateral breast cancer bore a 1.25-fold augmented risk of distant metastasis relative to unilateral cases [22]. Given the infrequency of bilateral breast cancer, our metastatic incidence results necessitate judicious interpretation.

For all metastatic patterns, we identified that patients with triple-negative breast cancer (TNBC) manifested the briefest survival span, whereas those with HR+/HER2+ exhibited the most favorable prognosis. Notably, our study pioneers the observation that metastatic site exhibited negligible impact on the survival of TNBC patients but displayed pronounced significance in HR+/HER2-. There was a discernible decrement in survival with each successive age decade, which can be partially attributed to the fact that older individuals might be predisposed to noncancerous mortality causes. Conversely, younger patients might be afforded enhanced access to locoregional therapeutic modalities, encompassing surgery, radiotherapy, and chemotherapy, thereby amplifying their survival prospects [23, 24]. Interestingly, our analysis revealed that patients with HR+/HER2+ demonstrated superior survival relative to the HR+/HER2- subtype, an outcome potentially influenced by the availability of more precision-targeted therapeutic regimens. In addition to this, recent studies have provided further insights into the molecular mechanisms driving metastatic spread in specific organs. Ganesan et al. [25] demonstrated that the EGFR-mediated PI3K/Akt/mTOR signaling pathway plays a crucial role in promoting lung metastasis in TNBC, highlighting the potential therapeutic targets for lung metastasis, a site more common in older patients in our study. Similarly, Zhang et al. [26] identified as a key regulator of liver metastasis, underscoring the importance of circRNAs in regulating liver metastasis in breast cancer. This is particularly relevant given the increased incidence of liver metastasis in older patients, as observed in our study. Furthermore, Xie et al. [27] utilized single-cell RNA sequencing to map the brain metastasis ecosystem, pinpointing as a promising therapeutic target. Brain metastasis, which was not significantly age-dependent in our

study, might benefit from therapies targeting these molecular markers.

This investigation is not devoid of limitations. (1) Screening modalities: Given that breast cancer screening guidelines don't invariably advocate for the employment of the most sensitive screening modalities, it's plausible that our estimates under-shoot the actual metastatic rates. (2) Data completeness: The SEER database does not furnish information on residence type, educational attainment, median household income, comorbidities, performance status, smoking habits, psychological health, or pivotal genomic data. These missing factors could influence the generalizability of our findings, especially when considering the diversity of factors that affect cancer outcomes. (3) Demographic representation: The SEER database includes data from 18 cancer registries in the United States, covering approximately 28% of the US population. Although this provides valuable insights, it may not fully represent the cancer characteristics of different demographic groups, especially those from non-US populations or under-represented groups. Additionally, the database relies on self-reported demographic information, which could introduce reporting biases, particularly with regard to race and marital status. (4) Metastatic details and treatment information: The SEER database does not delineate the precise metastatic locations, metastatic count, or exhaustive treatment details. As a result, we were unable to fully explore the complexity of multisite metastasis or the impact of specific therapies on metastatic spread. (5) De novo metastasis data: This study exclusively hinged on de novo metastasis data, eschewing data on relapsed metastatic cancer, and relying solely on overall survival metrics. The exclusion of relapsed cancer data may limit the applicability of our findings to patients who experience later-stage metastasis following initial treatment. (6) Retrospective design and data limitations: As a retrospective cohort study, this investigation has inherent limitations, including the inability to establish causal relationships and the potential for bias in data collection. Although the SEER database provides valuable population-level data, its retrospective nature means that the data were not collected with a specific research hypothesis in mind, which may limit the ability to control for all potential confounding factors. Furthermore, the lack of detailed clinical data (such as treatment regimens, performance status, or comorbidities) means that the influence of certain confounders on metastatic progression and survival outcomes may not be fully captured. These limitations could impact the generalizability of our findings, particularly in more specific patient subgroups, such as those with complex clinical conditions or those receiving specific treatments not recorded in the SEER dataset.

The variation in metastatic patterns among different molecular subtypes, such as HR+/HER2+ and HR-/HER2-, contributes critical insights into treatment decisions. Patients with HR+/HER2+ breast cancer, for example, who are more likely to develop bone and lung metastases, may benefit from combination therapies targeting both HER2 receptors and hormonal pathways. In contrast, those with HR-/HER2- subtypes, who are more prone to lung and liver metastases, may require different systemic therapies. These findings suggest the need for a more patient-centered approach to treatment, where the clinical team, including oncologists,

radiologists, and surgeons, collaborate to develop individualized treatment plans that integrate both systemic and localized therapies based on the patient's specific metastatic profile. These personalized strategies are not only likely to improve clinical outcomes but could also help make cancer care more cost-effective. By focusing on high-risk groups, healthcare systems can allocate resources more efficiently, avoiding overuse of expensive treatments in lower risk populations and improving the management of patients who need more intensive interventions.

Based on the findings of this study, several important directions for future research are warranted. (1) Multicenter, prospective studies: Our study, being retrospective in nature, lays the foundation for further prospective research. A multicenter, longitudinal study could help confirm and expand our findings, particularly in populations outside the United States and those underrepresented in SEER. This would also allow for a more detailed exploration of how different geographic regions or healthcare systems influence metastatic patterns and survival outcomes. (2) Molecular biomarkers and genomic studies: Given the role of biological factors in metastatic spread, future studies should focus on identifying specific molecular biomarkers that contribute to metastasis. Exploring the genetic and epigenetic landscape of breast cancer could provide critical insights into the molecular mechanisms underlying metastatic dissemination and help identify potential targets for therapeutic intervention. (3) Relapsed metastatic disease: Our study focused exclusively on de novo metastatic cases. Future research should include relapsed metastatic disease, as this patient group may exhibit distinct metastatic behaviors and survival outcomes. Investigating the differences between primary and recurrent metastasis could inform tailored treatment strategies. (4) Impact of specific therapies on metastasis: Further research should explore how specific treatments (e.g., chemotherapy, radiotherapy, and targeted therapies) influence metastatic patterns. This could involve examining the effect of treatments on specific metastatic sites (such as bone or brain) and evaluating the long-term outcomes of these therapies in preventing or controlling metastasis. (5) Age-related differences in metastatic patterns: Given the age-related differences observed in metastatic patterns, future studies could investigate the underlying mechanisms driving these differences, including immune system alterations and hormonal influences. Specifically, studies comparing metastatic behavior in younger versus older patients could provide valuable insights into how age affects metastatic dissemination and response to treatment.

## 5 | Conclusions

This study provides meaningful insights into the heterogeneous metastatic patterns of breast cancer, emphasizing the influence of age and molecular subtype on the distribution and prognosis of metastatic disease. Our findings highlight the necessity of adopting a more personalized management to both screening and treatment. Stratified screening strategies—such as prioritizing bone surveillance in younger patients and focusing on lung and liver assessments in older patients—may facilitate earlier detection and more effective interventions. Moreover, treatment plans should be tailored according to the specific metastatic profiles and tumor biology of each patient.

Personalized therapeutic strategies, including the use of bone-targeted agents or localized therapies for brain metastases, can significantly improve clinical outcomes. These findings support the implementation of age- and subtype-specific protocols in managing metastatic breast cancer to optimize patient care and healthcare resource utilization.

### Author Contributions

**Xiangyi Kong:** conceptualization (lead), investigation (lead), writing – original draft (equal). **Qiang Liu:** data curation (equal), formal analysis (lead), methodology (lead). **Zheng Qu:** investigation (equal), writing – original draft (equal), writing – review and editing (equal). **Xiangyu Wang:** investigation (equal), project administration (equal), software (equal). **Wenxiang Zhang:** methodology (equal), resources (equal), visualization (equal). **Yulu Liu:** conceptualization (equal), project administration (equal), software (equal). **Robert Coleman:** conceptualization (equal), supervision (equal), visualization (equal). **Chunqing Lin:** supervision (equal), validation (equal). **Jing Wang:** funding acquisition (equal), supervision (lead), validation (lead).

### Acknowledgments

The present study was conducted in compliance with the National Cancer Institute's SEER limited-use data agreement.

### Ethics Statement

The analysis followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines. Ethical review and approval were waived for this study because the SEER Program provides secondary and open databases.

In accordance with the Declaration of Helsinki, the research was registered in Prospero (<https://www.crd.york.ac.uk/prospero/>).

### Consent

Informed consent was waived due to the use of de-identified data.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Data Availability Statement

The dataset supporting the conclusions of this article is available in the SEER repository.

### References

1. Breast Cancer Statistics in Australia National Breast Cancer Foundation, <https://nbcf.org.au/about-breast-cancer/breast-cancer-stats/>.
2. A. J. Redig and S. S. McAllister, "Breast Cancer as a Systemic Disease: A View of Metastasis," *Journal of Internal Medicine* 274, no. 2 (2013): 113–126, <https://doi.org/10.1111/joim.12084>.
3. H. Kennecke, R. Yerushalmi, R. Woods, et al., "Metastatic Behavior of Breast Cancer Subtypes," *Journal of Clinical Oncology* 28, no. 20 (2010): 3271–3277, <https://doi.org/10.1200/JCO.2009.25.9820>.
4. H. Wang, C. Zhang, J. Zhang, L. Kong, H. Zhu, and J. Yu, "The Prognosis Analysis of Different Metastasis Pattern in Patients With Different Breast Cancer Subtypes: A SEER Based Study," *Oncotarget* 8, no. 16 (2017): 26368–26379, <https://doi.org/10.18632/oncotarget.14300>.
5. W. Xiao, S. Zheng, A. Yang, et al., "Breast Cancer Subtypes and the Risk of Distant Metastasis at Initial Diagnosis: A Population-Based

- Study," *Cancer Management and Research* 10 (2018): 5329–5338, <https://doi.org/10.2147/cmar.s176763>.
6. R. Wang, Y. Zhu, X. Liu, X. Liao, J. He, and L. Niu, "The Clinicopathological Features and Survival Outcomes of Patients With Different Metastatic Sites in Stage IV Breast Cancer," *BMC Cancer* 19, no. 1 (2019): 1091, <https://doi.org/10.1186/s12885-019-6311-z>.
7. C. G. Yedjou, J. N. Sims, L. Miele, et al., "Health and Racial Disparity in Breast Cancer," *Advances in Experimental Medicine and Biology* 1152 (2019): 31–49, [https://doi.org/10.1007/978-3-030-20301-6\\_3](https://doi.org/10.1007/978-3-030-20301-6_3).
8. Y. Liu, D. Wang, Z. Yang, et al., "Marital Status Is an Independent Prognostic Factor in Inflammatory Breast Cancer Patients: An Analysis of the Surveillance, Epidemiology, and End Results Database," *Breast Cancer Research and Treatment* 178, no. 2 (2019): 379–388, <https://doi.org/10.1007/s10549-019-05385-8>.
9. M. Watson, M. Saraiya, F. Ahmed, et al., "Using Population-Based Cancer Registry Data to Assess the Burden of Human Papillomavirus-Associated Cancers in the United States: Overview of Methods," *Cancer* 113, no. S10 (2008): 2841–2854, <https://doi.org/10.1002/cncr.23758>.
10. J. Snaedal, "The Helsinki Declaration," [in Icelandic], *Læknablaðið* 100, no. 3 (2014): 135, <https://doi.org/10.17992/lbl.2014.03.533>.
11. K. Asplund and G. Hermerén, "The Need to Revise the Helsinki Declaration," *Lancet* 389, no. 10075 (2017): 1190–1191, [https://doi.org/10.1016/s0140-6736\(17\)30776-6](https://doi.org/10.1016/s0140-6736(17)30776-6).
12. M. Czarkowski, "Kolejna Nowelizacja Deklaracji Helsińskiej [Helsinki Declaration–Next Version]," [in Polish], *Polski Merkurusz Lekarski* 36, no. 215 (2014): 295–297.
13. E. von Elm, D. G. Altman, M. Egger, S. J. Pocock, P. C. Gøtzsche, and J. P. Vandenbroucke, "The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies," *Annals of Internal Medicine* 147, no. 8 (2007): 573–577, <https://doi.org/10.7326/0003-4819-147-8-200710160-00010>.
14. G. Mathew and R. Agha, "STROCSS 2021: Strengthening the Reporting of Cohort, Cross-Sectional and Case-Control Studies in Surgery," *International Journal of Surgery Open* 37 (2021): 100430, <https://doi.org/10.1016/j.ijso.2021.100430>.
15. G. Farshid and D. Walters, "Molecular Subtypes of Screen-Detected Breast Cancer," *Breast Cancer Research and Treatment* 172, no. 1 (2018): 191–199, <https://doi.org/10.1007/s10549-018-4899-3>.
16. J. Xie, Y.-Y. Ying, B. Xu, Y. Li, X. Zhang, and C. Li, "Metastasis Pattern and Prognosis of Male Breast Cancer Patients in Us: A Population-Based Study From SEER Database," *Therapeutic Advances in Medical Oncology* 11 (2019): 1758835919889003, <https://doi.org/10.1177/1758835919889003>.
17. A. M. Martin, D. N. Cagney, P. J. Catalano, et al., "Brain Metastases in Newly Diagnosed Breast Cancer," *JAMA Oncology* 3, no. 8 (2017): 1069, <https://doi.org/10.1001/jamaoncol.2017.0001>.
18. W. Fang, Y. Huang, X. Han, J. Peng, and M. Zheng, "Characteristics of Metastasis and Survival Between Male and Female Breast Cancer With Different Molecular Subtypes: A Population-Based Observational Study," *Cancer Medicine* 11, no. 3 (2022): 764–777, <https://doi.org/10.1002/cam4.4469>.
19. R. De Luca, P. Dorangricchia, L. Salerno, G. Lo Coco, and G. Cicero, "The Role of Couples' Attachment Styles in Patients' Adjustment to Cancer," *Oncology* 92, no. 6 (2017): 325–334, <https://doi.org/10.1159/000455956>.
20. A. E. Kanters, A. M. Morris, P. H. Abrahamse, L. Mody, and P. A. Suwanabol, "The Effect of Peer Support on Colorectal Cancer Patients' Adherence to Guideline-Concordant Multidisciplinary Care," *Diseases of the Colon and Rectum* 61, no. 7 (2018): 817–823, <https://doi.org/10.1097/dcr.0000000000001067>.
21. N. Coburn, J. Fulton, D. N. Pearlman, C. Law, B. DiPaolo, and B. Cady, "Treatment Variation by Insurance Status for Breast Cancer Patients," *Breast Journal* 14, no. 2 (2008): 128–134, <https://doi.org/10.1111/j.1524-4741.2007.00542.x>.
22. D. E. Heron, L. T. Komarnicky, T. Hyslop, G. F. Schwartz, and C. M. Mansfield, "Bilateral Breast Carcinoma: Risk Factors and Outcomes for Patients With Synchronous and Metachronous Disease," *Cancer* 88, no. 12 (2000): 2739–2750.
23. R. Warschkow, U. Güller, I. Tarantino, et al., "Improved Survival After Primary Tumor Surgery in Metastatic Breast Cancer: A Propensity-Adjusted, Population-Based SEER Trend Analysis," *Annals of Surgery* 263, no. 6 (2016): 1188–1198, <https://doi.org/10.1097/SLA.0000000000001302>.
24. A. A. Sofi, I. Mohamed, M. Koumaya, and Z. Kamaluddin, "Local Therapy in Metastatic Breast Cancer Is Associated With Improved Survival," *American Journal of Therapeutics* 20, no. 5 (2013): 487–492, <https://doi.org/10.1097/MJT.0b013e31822119c5>.
25. K. Ganesan, C. Xu, J. Wu, et al., "Ononin Inhibits Triple-Negative Breast Cancer Lung Metastasis by Targeting the EGFR-Mediated PI3K/Akt/mTOR Pathway," *Science China Life Sciences* 67, no. 9 (2024): 1849–1866, <https://doi.org/10.1007/s11427-023-2499-2>.
26. Y. Zhang, Y. Tan, J. Yuan, et al., "CircLIFR-007 Reduces Liver Metastasis via Promoting hnRNPA1 Nuclear Export and YAP Phosphorylation in Breast Cancer," *Cancer Letters* 592 (2024): 216907, <https://doi.org/10.1016/j.canlet.2024.216907>.
27. J. Xie, A. Yang, Q. Liu, et al., "Single-Cell RNA Sequencing Elucidated the Landscape of Breast Cancer Brain Metastases and Identified ILF2 as a Potential Therapeutic Target," *Cell Proliferation* 57, no. 11 (2024): e13697, <https://doi.org/10.1111/cpr.13697>.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section.