Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Reproductive and Family-Structure Factors and Tumor Characteristics at Diagnosis in Iranian Women with Breast Cancer: A Retrospective Single-Center Study

Fereshteh Talebi
Shahid Beheshti University of Medical Sciences
Elham Ehsani-Chimeh
Tehran University of Medical Sciences
Tannaz Malakouti
ttmalakouti1990@gmail.com
Shahid Beheshti University of Medical Sciences
Ali Motlagh
Shahid Beheshti University of Medical Sciences

Research Article

Keywords: Breast Neoplasms, Reproductive History, Ki-67 Antigen, Tumor Biomarkers, Breast Cancer

Posted Date: March 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8720147/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Breast cancer is the most frequently diagnosed malignancy among women worldwide. Tumor characteristics at diagnosis vary considerably and may be shaped by biological, reproductive, and demographic factors. This study assessed whether reproductive and familial history are independently associated with tumor characteristics, with emphasis on proliferative activity at diagnosis.

Methods

We conducted a retrospective single-center study of 321 women diagnosed with breast cancer between 2017 and 2020 at Imam Hossein Hospital in Tehran, Iran. Tumor-related variables were abstracted from medical records, and reproductive and familial history variables were obtained from medical records when available and supplemented by structured telephone interviews when necessary. Associations between reproductive and familial factors and tumor characteristics (using Ki67 as the primary outcome) were evaluated using multivariable regression models appropriate to outcome type.

Results

Age was consistently associated with Ki67 and with prespecified secondary outcomes, including tumor grade, tumor size category, and metastatic disease at diagnosis. In adjusted analyses, each additional year of age was associated with an approximate 0.4 percentage-point decrease in Ki67, higher odds of metastatic disease at diagnosis, and lower odds of grade 3 tumors. No reproductive or familial history variables showed consistent independent associations with tumor characteristics after multivariable adjustment.

Conclusion

Age emerged as the primary factor related to tumor proliferative activity and selected tumor characteristics at diagnosis, whereas reproductive and familial history variables showed no independent associations after adjustment. Overall, the findings emphasize the importance of age-related influences in breast cancer presentation and support further investigation in larger populations to inform age-responsive risk assessment and early detection strategies.

Introduction

Breast cancer (BC) is the most commonly diagnosed malignancy among women worldwide (1). According to GLOBOCAN 2022, breast cancer accounted for approximately 2.3 million new cases and 666,000 deaths globally (1). Understanding factors associated not only with breast cancer risk but also with tumor characteristics at diagnosis remains central to cancer control efforts, as tumor phenotype at presentation is clinically important and varies across clinically relevant tumor characteristics (2, 3).

Epidemiologic studies have identified numerous factors associated with breast cancer risk, including lifestyle-related exposures, reproductive and hormonal factors, and non-modifiable characteristics such as genetic background and family history (3–5). However, risk factors for disease occurrence do not necessarily correspond to determinants of tumor biology or tumor characteristics at diagnosis. In Iran, a systematic review reported associations between several reproductive and lifestyle factors and breast cancer risk, while breastfeeding, physical activity, and selected dietary patterns were associated with lower risk (6). Whether similar factors are independently associated with tumor characteristics at diagnosis remains less well defined.

Reproductive and familial characteristics have also been implicated in shaping breast cancer biology. Parity and breastfeeding have been associated with differences in tumor subtype distribution and, in some studies, with lower prevalence of aggressive tumor phenotypes (7, 8). Familial factors such as birth order and parental age have been explored primarily in relation to breast cancer risk, whereas their potential associations with tumor characteristics at presentation are less well understood and findings have been inconsistent across populations (9).

Tumor biomarkers and histopathologic features play a central role in breast cancer classification, prognosis, and treatment selection. Estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status are routinely assessed at diagnosis and provide critical prognostic and predictive information(10–14). Ki67, a marker of cellular proliferation, is widely used to assess tumor growth fraction and aggressiveness. Higher Ki67 levels are generally associated with poorer prognosis, although its clinical interpretation may vary according to tumor subtype and treatment context(15).

Given the heterogeneity of breast cancer and ongoing uncertainty regarding the independent contribution of reproductive and familial factors to tumor phenotype at diagnosis, we conducted a retrospective study among Iranian women with breast cancer. The primary objective was to evaluate whether reproductive and familial factors are independently associated with tumor proliferative activity, measured by Ki67, after adjustment for age. Secondary objectives were to examine associations with tumor grade and metastatic disease at diagnosis.

Materials and Methods

Study Design

This was a retrospective, single-center observational study of women diagnosed with breast cancer between 2017 and 2020 at Imam Hossein Hospital in Tehran, Iran. The study evaluated associations between selected reproductive and familial factors and tumor characteristics at diagnosis.

Population and Data Collection

A total of 321 women with a diagnosis of breast cancer were included in this retrospective cross-sectional study. Analyses were conducted using a complete-case approach for each outcome; therefore, analytic sample sizes varied across models depending on outcome availability.

Tumor-related variables at diagnosis, including TNM stage (AJCC 8th edition), estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) status, tumor grade, and Ki67 percentage, were abstracted exclusively from pathology and clinical records.

Reproductive and familial factors, including parity, breastfeeding history, parental age, birth order, and pregnancy intervals, were obtained from medical records when available and were supplemented by structured telephone interviews when necessary.

The interview questionnaire, which served as the data collection tool, was specifically developed for this study based on a review of scientific evidence and was not adapted from any previously published instrument. An English version of the questionnaire is provided as a supplementary file (Supplementary Material 1).

Given the retrospective design and partial reliance on self-reported reproductive history data, the potential recall bias was considered in the interpretation of the results.

Statistical Analysis

Statistical analyses were performed using SPSS version 20 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean (standard deviation) and median (range), and categorical variables as frequency (percentage).

The primary outcome was Ki67 (%) at diagnosis, analyzed as a continuous variable using linear regression. Secondary outcomes were tumor grade (grade 3 vs grades 1–2) and metastatic disease at diagnosis (present vs absent), analyzed using binary logistic regression. Analyses of receptor status (ER, PR, HER2), tumor size category, and nodal status category were considered exploratory and are reported in the Supplementary Material; ordinal logistic regression was used for ordered categorical outcomes (tumor size and nodal status).

Multivariable models adjusted for age and a prespecified set of reproductive and familial covariates selected a priori based on clinical relevance, including number of siblings, number of children, breastfeeding duration during infancy, cumulative breastfeeding duration for own children, parental age, age at first pregnancy, maternal pregnancy interval, patient pregnancy interval, and birth order. Analyses were conducted using complete-case analysis for each outcome; participants with missing outcome data were excluded from the corresponding models.

All tests were two-sided. Results are reported as β coefficients (linear regression) or odds ratios (logistic regression) with 95% confidence intervals. Given the prespecified outcome hierarchy and the exploratory nature of secondary and exploratory analyses, emphasis was placed on effect estimates and confidence intervals rather than formal adjustment for multiple comparisons.

Model assumptions were assessed. For linear regression models, linearity was evaluated using residual-versus-fitted plots, normality of residuals using Q-Q plots, and homoscedasticity by inspection of standardized residuals. For ordinal logistic regression models, the proportional odds assumption was assessed using the Brant test and was satisfied for all fitted models.

This study is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for observational studies(16).

Results

Baseline characteristics

The analysis included 321 women diagnosed with breast cancer, with a mean age of 52.5 years (SD, 11.5). At diagnosis, approximately one quarter of patients (25.9%) had metastatic disease. The distribution of tumor grade and biomarker characteristics reflected a predominance of intermediate-grade tumors, with wide variability in tumor proliferative activity as measured by Ki67. Baseline demographic, reproductive, and clinicopathologic characteristics of the study population are summarized in Table 1. A small proportion of tumor-related variables had unknown values and were excluded from the corresponding regression analyses.

Table 1. Baseline demographic, reproductive, and clinicopathologic characteristics of the study population (n=321)

Continuous variables		
Variable	**Mean (SD)**	**Median (Range)**
Age, years	52.5 (11.5)	52 (25–87)
Number of siblings	5.4 (2.6)	5 (1–16)
Number of children	2.7 (1.5)	2 (1–10)
Duration of breastfeeding during infancy (patient as infant), years	1.8 (0.6)	2 (0–3)
Cumulative duration of breastfeeding (own children), years	3.9 (2.6)	4 (0–16)
Mother's age, years	26.0 (7.7)	25 (12–53)
Father's age, years	33.4 (8.9)	32 (16–60)
Age at first pregnancy, years	22.7 (5.6)	22 (10–47)
Mother's pregnancy interval, years	13.7 (7.0)	13 (1–40)
Patient pregnancy interval, years	9.8 (6.3)	9 (2–35)
Birth order	3.3 (2.2)	3 (1–13)
Ki67, %	26 (18)	20 (2–95)
Tumor characteristics		
Variable	**n (%)**	
Histologic grade		
Grade 1	17 (5.3)	
Grade 2	238 (74.1)	
Grade 3	45 (14.0)	
Unknown	21 (6.5)*	
ER status		
Negative	75 (23.4)	
Positive	246 (76.6)	
PR status		
Negative	99 (30.8)	
Positive	222 (69.2)	
HER2 status		
Negative	216 (67.3)	
Positive	105 (32.7)	
Nodal status		
N0	120 (37.4)	
N1	121 (37.7)	
N2	54 (16.8)	
N3	24 (7.5)	
Unknown	2 (0.6)*	
Tumor size		
T1	66 (20.6)	
T2	203 (63.2)	
T3	41 (12.8)	
T4	8 (2.5)	
Unknown	3 (0.9)*	
Metastatic disease at diagnosis		
Absent	236 (73.5)	

Present	83 (25.9)
Unknown	2 (0.6)*

• Duration of breastfeeding during infancy refers to the reported duration of breastfeeding when the patient was an infant.

• Mother's pregnancy interval refers to the interval between the mother's pregnancies preceding and following the index patient.

• Patient pregnancy interval was defined as the time (in years) between the first and last pregnancy.• Percentages are calculated based on the total study population unless otherwise indicated.

*• *Categories with unknown values were excluded from regression analyses.*

Ki67 (%)

In univariable linear regression analyses, increasing age and longer cumulative duration of breastfeeding were associated with lower Ki67 percentage. In multivariable models adjusting for age and prespecified reproductive and familial covariates, increasing age remained independently associated with lower tumor proliferative activity, corresponding to an approximate 0.4 percentage-point decrease in Ki67 per additional year of age (Table 2). In contrast, the association between cumulative breastfeeding duration and Ki67 was attenuated and did not persist after adjustment. No other reproductive or familial history variables showed evidence of independent associations with Ki67 after adjustment.

Table 2. Univariable and multivariable linear regression analyses for factors associated with Ki67 (%) at diagnosis

Variable	Univariable β (95% CI)	p-value	Multivariable β (95% CI)	p-value
Age, years	-0.003 (-0.004 to -0.001)	0.004	-0.004 (-0.007 to -0.002)	0.001
Number of siblings	0.002 (-0.005 to 0.010)	0.54	—	—
Number of children	-0.002 (-0.020 to 0.010)	0.72	—	—
Duration of breastfeeding during infancy (patient as infant), years	-0.020 (-0.060 to 0.020)	0.29	—	—
Cumulative duration of breastfeeding (own children), years	-0.010 (-0.020 to -0.001)	0.03	-0.006 (-0.020 to 0.010)	0.37
Mother's age, years	0.002 (-0.001 to 0.005)	0.24	0.001 (-0.003 to 0.004)	0.86
Father's age, years	0.001 (-0.001 to 0.004)	0.31	—	—
Age at first pregnancy, years	-0.001 (-0.005 to 0.003)	0.52	—	—
Mother's pregnancy interval, years	-0.002 (-0.006 to 0.001)	0.17	-0.002 (-0.006 to 0.001)	0.20
Patient pregnancy interval, years	-0.0005 (-0.004 to 0.003)	0.78	—	—
Birth order	0.007 (-0.003 to 0.020)	0.17	0.010 (-0.002 to 0.020)	0.10

- *β coefficients represent the change in Ki67 percentage per one-unit increase in the predictor.*
- *Multivariable models adjusted for age and a prespecified set of reproductive and familial covariates selected a priori based on clinical relevance.*
- *A negative β coefficient indicates lower Ki67 values with increasing predictor levels.*
- *"—" indicates variables not retained in the multivariable model.*
- *n = 321 unless otherwise indicated*

ER, PR, and HER2 status

Exploratory analyses are presented in Supplementary Table S1. While several variables showed statistically significant associations with receptor status in univariable models, these associations did not persist after multivariable adjustment, and no variable demonstrated a consistent independent association across models.

Tumor size, nodal status, metastasis, and tumor grade

In univariable analyses, increasing age was associated with tumor size category. In multivariable ordinal logistic regression analyses adjusting for age and prespecified reproductive and familial covariates, increasing age remained independently associated with tumor size, with each additional year of age corresponding to approximately 4% lower odds of being in a higher (more advanced) tumor size category (Table 3).

No demographic, reproductive, or familial history variables were independently associated with nodal status in univariable or multivariable ordinal logistic regression analyses.

With respect to metastatic disease at diagnosis, no variables showed a statistically significant association with metastatic disease in univariable analyses. In multivariable logistic regression analysis, increasing age was independently associated with metastatic disease, with each additional year of age corresponding to approximately 3% higher odds of metastasis at diagnosis (Table 3).

Similarly, no reproductive or familial history variables were associated with tumor grade in univariable analyses. In multivariable logistic regression analysis, increasing age was independently associated with tumor grade, with each additional year of age corresponding to approximately 3% lower odds of grade 3 disease compared with grades 1–2 (Table 3).

Table 3. Univariable and multivariable regression analyses of age and reproductive/familial factors in relation to tumor size, nodal status, metastatic disease, and tumor grade at diagnosis

Variable	Univariable OR (95% CI) Tumor Size	p	Multivariable OR (95% CI) Tumor Size	p	Univariable OR (95% CI) Nodal Status	p	Multivariable OR (95% CI) Nodal Status	p	Univariable OR (95% CI) Metastasis	p	Multivariable OR (95% CI) Metastasis	p
Age, years	0.97 (0.95–0.99)	0.04	0.96 (0.93–0.99)	0.02	0.99 (0.97–1.01)	0.59	—	—	1.02 (1.00–1.04)	0.11	1.03 (1.01–1.05)	0.04
Number of siblings	1.03 (0.94–1.13)	0.50	—	—	0.98 (0.90–1.06)	0.65	—	—	1.04 (0.94–1.15)	0.43	—	—
Number of children	1.04 (0.90–1.22)	0.06	1.11 (0.87–1.42)	0.41	0.93 (0.81–1.08)	0.37	0.95 (0.77–1.16)	0.62	1.06 (0.89–1.26)	0.51	—	—
Duration of breastfeeding during infancy (patient as infant), years	1.16 (0.78–1.72)	0.47	—	—	1.01 (0.70–1.46)	0.96	—	—	1.47 (0.89–2.41)	0.13	1.28 (0.77–2.14)	0.34
Cumulative duration of breastfeeding (own children), years	1.01 (0.92–1.10)	0.89	—	—	0.93 (0.85–1.01)	0.10	0.96 (0.85–1.09)	0.53	1.05 (0.95–1.16)	0.31	—	—
Mother's age, years	1.00 (0.97–1.03)	0.87	—	—	0.99 (0.96–1.02)	0.44	1.02 (0.97–1.07)	0.43	0.99 (0.96–1.03)	0.71	—	—
Father's age, years	0.99 (0.96–1.02)	0.07	0.97 (0.93–1.00)	0.06	0.99 (0.96–1.01)	0.29	0.97 (0.93–1.01)	0.13	1.01 (0.98–1.04)	0.51	—	—
Age at first pregnancy, years	1.00 (0.96–1.05)	0.57	—	—	1.01 (0.97–1.05)	0.72	—	—	1.02 (0.97–1.06)	0.51	—	—
Mother's pregnancy interval, years	1.00 (0.96–1.03)	0.83	—	—	1.00 (0.96–1.03)	0.80	—	—	1.03 (0.99–1.07)	0.14	1.02 (0.98–1.07)	0.29
Patient pregnancy interval, years	0.97 (0.93–1.02)	0.22	0.99 (0.93–1.04)	0.67	1.01 (0.97–1.05)	0.51	—	—	1.02 (0.97–1.07)	0.41	—	—
Birth order	1.07 (0.96–1.19)	0.20	1.10 (0.96–1.26)	0.40	0.97 (0.88–1.06)	0.47	—	—	0.94 (0.83–1.06)	0.31	—	—

- *Tumor size and nodal status were analyzed using ordinal logistic regression (AJCC categories).*
- *Metastasis at diagnosis and tumor grade were analyzed using binary logistic regression (tumor grade: grade 3 vs grades 1–2).*
- *Odds ratios >1 indicate higher odds of more advanced disease or higher category, depending on the outcome.*
- *Multivariable models were adjusted for age and selected reproductive and familial covariates identified a priori.*
- *"—" indicates variables not retained in the multivariable model.*

Sample sizes vary due to exclusion of cases with missing outcome data.

Discussion

Principal Findings

In this retrospective single-center study of 321 women with breast cancer, age at diagnosis emerged as the most consistent factor associated with tumor characteristics at presentation. Increasing age was independently associated with lower tumor proliferative activity (Ki67), lower odds of grade 3 disease, and lower odds of a more advanced tumor size category, while being associated with higher odds of metastatic disease at diagnosis. In contrast, reproductive and familial factors did not demonstrate consistent independent associations with tumor characteristics after multivariable adjustment.

Ki67

In this single-center study, increasing age at diagnosis was independently associated with lower tumor proliferative activity, as reflected by Ki67, even after adjustment for prespecified reproductive and familial covariates. This pattern is broadly concordant with the established observation that breast cancers arising at younger ages tend to display higher proliferation and other aggressive biological features, and with evidence supporting Ki67 as a clinically meaningful marker of tumor biology and prognosis in breast cancer (17). An alternative explanation that may contribute to the observed age-Ki67 gradient is compositional: age may correlate with unmeasured tumor subtype distributions (e.g., luminal versus non-luminal disease) and treatment or referral pathways, which could influence measured proliferation at diagnosis(18, 19). In our analyses, the univariable association between cumulative breastfeeding duration and lower Ki67 did not persist after adjustment, suggesting that any apparent relationship between breastfeeding history and proliferation in this dataset may be explained by age or other correlated reproductive factors rather than an independent effect.

Tumor grade

We additionally observed that increasing age was associated with lower odds of grade 3 disease (vs grades 1–2) in adjusted analyses. This finding aligns with literature describing age-related differences in tumor differentiation and subtype patterns, in which older age groups more commonly present with biologically indolent phenotypes(18, 19). At the same time, grade is an imperfect proxy for tumor aggressiveness and may capture different biological axes than proliferation (Ki67), which can vary substantially within the same grade category. Differences in grading practices, interobserver variability, and the relatively small number of grade 1 tumors in this study (necessitating a binary grade outcome) may also partly explain why the age–grade association observed here could differ in magnitude or direction from other settings.

Metastatic disease at diagnosis

In contrast to the inverse associations observed between age and both Ki67 and grade 3 disease, increasing age was independently associated with higher odds of metastatic disease at diagnosis in adjusted analyses. This divergence indicates that associations between age and tumor biological characteristics do not necessarily parallel associations with disease extent at presentation. Consistent with this finding, population-based data have reported an increasing proportion of metastatic breast cancer at diagnosis with advancing age, with substantially higher rates observed among the oldest age groups (20). In the present study, no reproductive or familial history variables were independently associated with metastatic disease, suggesting that these factors did not explain variation in metastatic presentation within this study. Because information on screening practices, diagnostic pathways, and timing of presentation was not available, the observed association between age and metastatic disease should be interpreted descriptively rather than causally.

Integrated interpretation and context

Taken together, our findings indicate that age is a consistent correlate of multiple dimensions of breast cancer presentation at diagnosis in this study, although not in a uniform direction. Increasing age was associated with lower tumor proliferative activity (Ki67) and lower odds of high-grade disease, while simultaneously being associated with higher odds of metastatic disease at presentation. In contrast, reproductive and familial history variables did not demonstrate consistent independent associations with tumor characteristics after multivariable adjustment. These patterns suggest that age captures both biological variation in tumor phenotype and non-biological influences related to disease detection and presentation.

The inverse associations between age and tumor proliferative activity or grade observed in this study are broadly concordant with prior epidemiologic and molecular literature describing age-related heterogeneity in breast cancer biology(18, 21). Previous studies have shown that established breast cancer risk factors may exhibit differential associations across clinically relevant tumor features rather than exerting uniform effects across all disease dimensions. Within this context, the absence of independent associations between reproductive or familial history variables and tumor characteristics after adjustment is consistent with reports suggesting that unadjusted associations may reflect confounding by age or correlated reproductive patterns rather than direct biological effects.

Importantly, the concurrent association of older age with lower proliferative and lower-grade tumors yet higher odds of metastatic disease underscores the distinction between intrinsic tumor biological features and disease extent at diagnosis. Tumor biology and stage at presentation, although related, represent partially independent dimensions of breast cancer characterization. Age-related differences in screening participation, diagnostic delay, comorbidity burden, or healthcare access (factors not captured in the present dataset) may contribute to the higher likelihood of metastatic disease among older patients despite less aggressive tumor biology. These findings emphasize the importance of interpreting age-related associations across tumor characteristics collectively rather than in isolation when evaluating observational data at diagnosis.

In contrast, reproductive and familial history variables did not demonstrate consistent independent associations with tumor characteristics after multivariable adjustment. Familial structure variables such as sibling number, birth order, and parental age also showed no independent associations, suggesting limited influence on tumor phenotype at diagnosis in this study. Although our findings indicated no direct associations between reproductive history or familial risk factors and tumor pathologic characteristics (e.g., grade, stage, or molecular phenotype), several epidemiologic studies have reported subtype-specific relationships. A large systematic review and meta-analysis found that later age at first birth, nulliparity/low parity, and breastfeeding were differentially associated with risks across luminal A/B, HER2-positive, and triple-negative breast cancers, with inconsistent strength and direction across studies (22). Similarly, another meta-analysis of 75 studies reported that reproductive factors—including age at menarche, age at first childbirth, parity,

breastfeeding, and age at menopause—were associated with varying risks across molecular subtypes, with stronger associations for luminal and HER2-positive tumors than triple-negative disease (23).

Additional observational evidence supports subtype-specific reproductive effects; for example, parity and age at first birth have been linked to different odds of triple-negative versus luminal A tumors in case-only analyses (24). These discrepancies with our results may be explained by methodological and contextual differences. First, studies vary in how reproductive and familial variables are defined and measured (e.g., continuous vs categorical age at first birth). Second, many prior analyses used unadjusted or minimally adjusted estimates, whereas our multivariable models controlled for age and other confounders, which may attenuate or eliminate apparent associations. Third, differences in population structure, genetic background, lifestyle, and access to health care across study settings can generate heterogeneous patterns of association, since reproductive behaviors and screening practices differ widely between populations.

Our findings, indicating that older age is associated with higher odds of metastatic disease at diagnosis, underscore the need for public health policies that go beyond general population-based screening and integrate risk assessment based on personal and familial history into the care pathway. In addition to regular mammography screening for the average-risk population, which has been shown to improve early detection and reduce mortality(25), the World Health Organization (WHO) emphasizes that early detection strategies should include increasing public awareness, improving access to diagnostic services, and strengthening rapid referral pathways to reduce diagnostic delays and improve clinical outcomes(26). For women at higher risk, targeted early screening and supplementary imaging are recommended; Established risk assessment tools, such as the National Cancer Institute Breast Cancer Risk Assessment Tool, as well as models like Gail, Tyrer-Cuzick, and BRCAPRO, can help identify women at elevated risk and determine who may benefit from earlier screening, shorter intervals, or supplemental imaging like MRI (27). Moreover, to reduce metastatic diagnoses in older individuals or populations with limited access to healthcare services, strengthening rapid diagnostic pathways is crucial, as healthcare systems with integrated diagnostic pathways have been shown to reduce diagnostic delays and lower the rates of advanced-stage disease (28). This multi-level approach, which includes general population screening, personal and familial risk assessment, targeted screening, and rapid referral pathways, can improve early detection, optimize healthcare resource allocation, and reduce the burden of late-stage diagnoses.

Methodological Context

This study employed a retrospective observational design with outcome-specific regression models and prespecified covariate adjustment to address potential confounding. Exploratory analyses were defined a priori, and all analyses were conducted in accordance with STROBE reporting guidelines.

Strengths and Limitations

The strengths of this study include detailed assessment of reproductive and familial history variables, comprehensive evaluation of tumor characteristics at diagnosis, and use of regression models appropriate to outcome type with explicit assessment of model assumptions.

However, several limitations warrant consideration.

The retrospective, single-center design limits generalizability.

Partial reliance on telephone interviews for reproductive history data introduces the potential for recall bias.

Information on screening practices, diagnostic delay, socioeconomic status, and healthcare access was not available, which may have influenced associations with stage-related outcomes such as metastatic disease at diagnosis.

Age at menarche was not available in our records and was not collected due to concerns about recall bias, although it has been associated with subtype-specific breast cancer risk in a recent systematic review and meta-analysis and with more aggressive tumor features in a recent observational study(23, 29).

Additionally, limited sample sizes in certain subgroups may have reduced statistical power to detect modest associations.

Conclusion

This study contributes to a more nuanced understanding of how patient age shapes breast cancer presentation, with important implications for epidemiologic research and cancer control strategies. Rather than supporting uniform effects of established reproductive or familial risk factors on tumor characteristics, our findings emphasize the relevance of tumor heterogeneity and the distinction between intrinsic tumor biology and disease extent at diagnosis. Within this framework, age likely reflects a combination of biological processes, clinical factors, and health system–related influences that extend beyond traditional risk profiles. From a public health and policy perspective, these results highlight the need to move beyond one-size-fits-all screening paradigms toward age-responsive and risk-informed early detection strategies, including timely diagnostic pathways and improved access to services for older populations. Integrating individualized risk assessment with population-level screening and strengthening diagnostic referral systems may help reduce late-stage presentation and improve equity in breast cancer outcomes. Future research linking life-course risk factors with healthcare access and system performance will be critical to informing more effective, targeted breast cancer prevention, and early detection policies.

Declarations

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Ethics Committee of Shahid Beheshti University of Medical Sciences, Tehran, Iran (approval number: IR.SBMU.MSP.REC.1398.255). Since the data were obtained from an educational hospital where written informed consent for the use of medical records is routinely secured from all patients upon admission, and given the retrospective design of the study utilizing existing clinical information, the Ethics Committee waived the requirement for additional informed consent. For participants who were contacted by telephone to complete missing reproductive and familial history data, verbal informed consent was obtained prior to the interview. This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. All data were handled confidentially and analyzed in de-identified form.

Consent for publication

Not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors thank the staff of Imam Hossein Hospital for assistance with accessing clinical and pathology records and the patients who participated in telephone interviews.

Use of artificial intelligence (AI) tools

Generative artificial intelligence (ChatGPT, OpenAI) was used solely to improve language clarity, grammar, and readability of the manuscript. The authors reviewed, edited, and verified all content and take full responsibility for the accuracy, integrity, and originality of the work. No AI tool was used to generate or analyze data, perform statistical analyses, or create results.

Author Contribution

F.T. was responsible for data collection and data acquisition and contributed to writing the original draft of the manuscript.T.M. wrote the original draft of the manuscript, performed data interpretation, revised the manuscript critically for important intellectual content, and managed the submission process as corresponding author.E.E.-C. provided scientific guidance, supervised manuscript development, and contributed to critical review and conceptual refinement of the study.A.M. contributed to study conception, provided senior supervision, and critically reviewed the manuscript.All authors read and approved the final manuscript.

Acknowledgement

The authors thank the staff of Imam Hossein Hospital for assistance with accessing clinical and pathology records and the patients who participated in telephone interviews.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to institutional policies and patient privacy considerations, but are available from the corresponding author on reasonable request, subject to ethics approval and data-sharing agreements as applicable.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2024;74(3):229–63.
2. Garcia-Closas M, Brinton LA, Lissowska J, Chatterjee N, Peplonska B, Anderson WF, et al. Established breast cancer risk factors by clinically important tumour characteristics. British Journal of Cancer. 2006;95(1):123–9.
3. Figueroa JD, Gierach GL, Duggan MA, Fan S, Pfeiffer RM, Wang Y, et al. Risk factors for breast cancer development by tumor characteristics among women with benign breast disease. Breast Cancer Res. 2021;23(1):34.
4. Łukasiewicz S, Czeczelewski M, Forma A, Baj J, Sitarz R, Stanisławek A. Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review. Cancers (Basel). 2021;13(17).
5. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer (Dove Med Press). 2019;11:151–64.
6. Shamshirian A, Heydari K, Shams Z, Aref AR, Shamshirian D, Tamtaji OR, et al. Breast cancer risk factors in Iran: a systematic review & meta-analysis. Hormone Molecular Biology and Clinical Investigation. 2020;41(4).
7. Figueroa JD, Davis Lynn BC, Edusei L, Titiloye N, Adjei E, Clegg-Lamptey J-N, et al. Reproductive factors and risk of breast cancer by tumour subtypes among Ghanaian women: A population-based case-control study. International Journal of Cancer. 2020;147(6):1535–47.

8. Fortner RT, Sisti J, Chai B, Collins LC, Rosner B, Hankinson SE, et al. Parity, breastfeeding, and breast cancer risk by hormone receptor status and molecular phenotype: results from the Nurses' Health Studies. Breast Cancer Res. 2019;21(1):40.
9. Nichols HB, Trentham-Dietz A, Sprague BL, Hampton JM, Titus-Ernstoff L, Newcomb PA. Effects of birth order and maternal age on breast cancer risk: modification by whether women had been breast-fed. Epidemiology. 2008;19(3):417–23.
10. Colomer R, González-Farré B, Ballesteros AI, Peg V, Bermejo B, Pérez-Mies B, et al. Biomarkers in breast cancer 2024: An updated consensus statement by the Spanish Society of Medical Oncology and the Spanish Society of Pathology. Clinical & Translational Oncology. 2024;26(12):2935–51.
11. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, et al. Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP clinical practice guideline update. Journal of Clinical Oncology. 2020;38(12):1346–66.
12. Hefti MM, Hu R, Knoblauch NW, Collins LC, Haibe-Kains B, Tamimi RM, et al. Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype. Breast Cancer Research. 2013;15(4):R68.
13. Li Y, Yang D, Yin X, Zhang X, Huang J, Wu Y, et al. Clinicopathological characteristics and breast cancer-specific survival of patients with single hormone receptor-positive breast cancer. JAMA Network Open. 2020;3(1):e1918160.
14. Marra A, Chandarlapaty S, Modi S. Management of advanced-stage HER2+ breast cancer: current evidence and future perspectives. Nature Reviews Clinical Oncology. 2024;21(3):185–202.
15. de Azambuja E, Cardoso F, de Castro Jr G, Colozza M, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12 155 patients. British Journal of Cancer. 2007;96:1504–13.
16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453–7.
17. Davey MG, Hynes SO, Kerin MJ, Miller N, Lowery AJ. Ki-67 as a Prognostic Biomarker in Invasive Breast Cancer. Cancers (Basel). 2021;13(17).
18. Mesa-Eguiagaray I, Wild SH, Williams LJ, Jin K, Bird SM, Brewster DH, et al. Breast cancer-specific survival by molecular subtype in different age groups of women in Scotland. Breast Cancer Research. 2025;27(1):59.
19. Anders CK, Hsu DS, Broadwater G, Acharya CR, Foekens JA, Zhang Y, et al. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol. 2008;26(20):3324–30.
20. Rottenberg Y, Naeim A, Uziely B, Peretz T, Jacobs JM. Breast cancer among older women: The influence of age and cancer stage on survival. Arch Gerontol Geriatr. 2018;76:60–4.
21. Chen HL, Zhou MQ, Tian W, Meng KX, He HF. Effect of Age on Breast Cancer Patient Prognoses: A Population-Based Study Using the SEER 18 Database. PLoS One. 2016;11(10):e0165409.
22. Lambertini M, Santoro L, Del Mastro L, Nguyen B, Livraghi L, Ugolini D, et al. Reproductive behaviors and risk of developing breast cancer according to tumor subtype: A systematic review and meta-analysis of epidemiological studies. Cancer Treat Rev. 2016;49:65–76.
23. Mao X, Omeogu C, Karanth S, Joshi A, Meernik C, Wilson L, et al. Association of reproductive risk factors and breast cancer molecular subtypes: a systematic review and meta-analysis. BMC Cancer. 2023;23(1):644.
24. Brouckaert O, Rudolph A, Laenen A, Keeman R, Bolla MK, Wang Q, et al. Reproductive profiles and risk of breast cancer subtypes: a multi-center case-only study. Breast Cancer Res. 2017;19(1):119.
25. Society AC. American Cancer Society recommendations for the early detection of breast cancer Atlanta (GA): American Cancer Society; 2023 [Available from: https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html.
26. Organization WH. Global breast cancer initiative implementation framework Geneva: World Health Organization; 2023 [updated 3 Feb 2023. Available from: https://www.who.int/publications/i/item/9789240067134https://www.who.int/initiatives/global-breast-cancer-initiative.
27. Institute NC. Breast Cancer Risk Assessment Tool (BCRAT) Bethesda (MD): National Cancer Institute; 2024 [updated 19 Nov 2024. Available from: https://bcrisktool.cancer.gov.
28. Richards MA, Westcombe AM, Love SB, Littlejohns P, Ramirez AJ. Influence of delay on survival in patients with breast cancer: a systematic review. Lancet. 1999;353(9159):1119–26.
29. Harris AR, Wang T, Heng YJ, Baker GM, Le PA, Wang J, et al. Association of early menarche with breast tumor molecular features and recurrence. Breast Cancer Res. 2024;26(1):102.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary8Feb.docx