SCIENTIFIC JOURNAL
CEREM-GO

DOI 10.37951/2675-5009.2026v7i20.202
ISSN 2675-5009
e26202
SCIENTIFIC ARTICLE - ORIGINAL

OVERALL SURVIVAL ACCORDING TO MOLECULAR SUBTYPES AND CLINICAL STAGING IN INVASIVE BREAST CARCINOMA PATIENTS: KAPLAN-MEIER ANALYSIS IN A HOSPITAL COHORT

JUAREZ ANTÔNIO DE SOUSA[1], WALDEMAR NAVES DO AMARAL[1], ANDRÉ MAROCCOLO DE SOUSA[2], DÉBORA NASCIMENTO DIAS NEVES[3], LAURA DE SOUSA E SOUSA[4], LUIZA DE SOUSA E SOUSA[4], ISABELLA RODRIGUES FERREIRA[5], PAMELA CHRISTINNY FERNANDES VIÊRA[5]

1. Doutor, Professor do Departamento de Ginecologia e Obstetrícia da Faculdade de Medicina da Universidade Federal de Goiás (UFG), Goiânia, GO, Brazil.
2. Médico Residente de Cirurgia Geral, Universidade de São Paulo (USP), São Paulo, SP, Brazil.
3. Médica, Universidade Federal de Goiás (UFG), Goiânia, GO, Brazil.
4. Acadêmica de Medicina, Universidade UniEvangélica, Anápolis, GO, Brazil.
5. Médica Residente de Ginecologia e Obstetrícia, Maternidade Aristina Cândida, Goiânia, GO, Brazil.

ABSTRACT

Introduction: Overall survival analysis according to molecular subtypes and clinical staging is essential for understanding prognosis in breast cancer. **Objective:** To estimate and compare overall survival of patients with invasive breast carcinoma according to clinical staging and main molecular subtypes in a hospital cohort from Central-West Brazil. **Methods:** Retrospective cohort study with 97 patients with confirmed invasive breast carcinoma, excluding ductal carcinoma in situ. Kaplan-Meier curves were compared by the log-rank test, stratified by staging (I–II vs. III–IV), triple-negative subtype, hormone receptor status, and HER2. Follow-up was estimated from the year of treatment initiation. **Results:** Mean estimated follow-up was 3.9 years (median: 3.0; range: 1–18 years). Six deaths (6.2%) and 9 recurrences (9.3%) were recorded. Advanced staging was associated with significantly worse survival (log-rank $p = 2.22 \times 10^{-10}$): death rate 50.0% vs. 1.1%. Triple-negative subtype showed worse survival vs. non-triple-negative (log-rank $p = 6.32 \times 10^{-4}$): 26.7% vs. 2.4% deaths. Hormone receptor negativity identified a higher-risk subgroup (log-rank $p = 6.27 \times 10^{-4}$): 22.7% vs. 1.3% deaths. HER2 status showed no significant difference (log-rank $p = 0.983$). **Conclusion:** Advanced staging, triple-negative subtype, and hormone receptor negativity were associated with worse overall survival. These findings reinforce the importance of molecular stratification and early diagnosis for prognostic individualization in breast oncology.

Keywords: Breast neoplasms, Survival analysis, Kaplan-Meier estimator, Molecular subtype, Prognosis.

INTRODUCTION

Survival analysis is a fundamental tool in prognostic research in oncology, allowing estimation of the probability of survival over time in populations with varying follow-up periods and incomplete events. In breast cancer, overall survival is determined by multiple clinical, pathological, and molecular factors, with clinical stage at diagnosis and the molecular profile of the tumor being the predictors with the greatest impact on the outcome.[1]

The molecular classification of breast cancer into subtypes — luminal A, luminal B, HER2-enriched, and triple-negative — has provided the basis for individualized therapeutic strategies and revolutionized the understanding of the biological behavior of the disease. The triple-negative subtype, due to the absence of hormone receptors and HER2 and the consequent lack of specific targeted therapies, has the poorest prognosis among the molecular subtypes. In contrast, hormone receptor-positive tumors exhibit less aggressive behavior and benefit from long-term adjuvant hormone therapy.[2]

In small hospital cohorts, the Kaplan–Meier method is particularly suitable for survival analysis because it does not impose parametric assumptions and naturally accommodates censored data. The log-rank test allows formal comparisons between subgroups, although with limited statistical power when the number of events is low — a limitation that should be explicitly considered when interpreting the results.[3]

The present study aimed to estimate and compare the overall survival of patients with invasive breast carcinoma according to clinical stage and the main molecular subtypes in a hospital cohort from the Central-West region of Brazil.[4]

METHODOLOGY

The Study Design and Population

A retrospective cohort study was conducted with 97 patients with a histologically confirmed diagnosis of invasive breast carcinoma treated at a private referral hospital for oncology in Goiás. Two cases of ductal carcinoma in situ (DCIS, stage 0) were excluded. The study was conducted in accordance with CNS Resolution No. 466/2012.

Definition of Follow-up Time

Follow-up time was estimated in years from the year treatment was initiated until 2025 (reference date), or until death when recorded. Patients without an event were treated as censored. This approach introduces imprecision in the estimation of exact follow-up times and is suitable for exploratory survival analyses using clinical service data.

Statistical Analysis

Overall survival curves were constructed using the Kaplan–Meier method and compared using the log-rank test, with stratification by: (1) clinical stage — early (I–II) vs. advanced (III–IV); (2) triple-negative subtype — present vs. absent; (3) hormone receptors — HR-positive (ER- or PR-positive) vs. HR-negative; and (4) HER2 status — positive vs. negative. All analyses were performed using R software (version 4.x) with the lifelines package. The significance level was set at 5%.

RESULTS

The Sample Characteristics and Follow-up

A total of 97 patients were included. The estimated mean follow-up was 3.9 years (median: 3.0 years; range: 1–18 years). Six deaths (6.2%) and 9 recurrences (9.3%) were recorded. Characteristics by subgroup are presented in Table 1, and the distribution of events is shown in Table 2.

Table 1. Sample Characteristics by Analysis Subgroup (n = 97).

Variable	n	%
Clinical stage		
Early (I–II)	87	89.7
Advanced (III–IV)	10	10.3
Molecular profile		
HR-positive	75	77.3
HR-negative	22	22.7
HER2-positive	19	19.6
HER2-negative	78	80.4
Triple-negative	15	15.5
Non-triple-negative	82	84.5
Estimated follow-up		
Mean (years)	3.9	—
Median (years)	3.0	—
Range (years)	1–18	—
Outcomes		
Death	6	6.2
Recurrence	9	9.3

HR: hormone receptors.

Table 2. Proportion of Deaths and Recurrences by Clinical and Molecular Subgroups (n = 97).

Subgroup	n	Deaths n (%)	Recurrences n (%))	p (log-rank)*
Clinical stage				
Early (I–II)	87	1 (1.1)	2 (2.3)	< 0.001
Advanced (III–IV)	10	5 (50.0)	7 (70.0)	< 0.001
Triple-negative subtype				
Non-triple-negative	82	2 (2.4)	5 (6.1)	< 0.001
Triple-negative	15	4 (26.7)	4 (26.7)	< 0.001
Hormone receptors				
HR-positive	75	1 (1.3)	2 (2.7)	< 0.001
HR-negative	22	5 (22.7)	7 (31.8)	< 0.001
HER2 status				
HER2-negative	78	5 (6.4)	5 (6.4)	0.983
HER2-positive	19	1 (5.3)	4 (21.1)	0.983

*p-value obtained using the log-rank test for the outcome of death between groups.

Figure 2. (A) Patient positioning for pre-oxygenation. (B) Video laryngoscopy.

Survival by Clinical Stage

Figure 1 presents the Kaplan–Meier curves stratified by clinical stage. The advanced-stage group (III–IV; n = 10) showed significantly lower overall survival than the early-stage group (I–II; n = 87), with a proportion of deaths of 50.0% vs. 1.1% (log-rank $p = 2.22 \times 10^{-10}$).

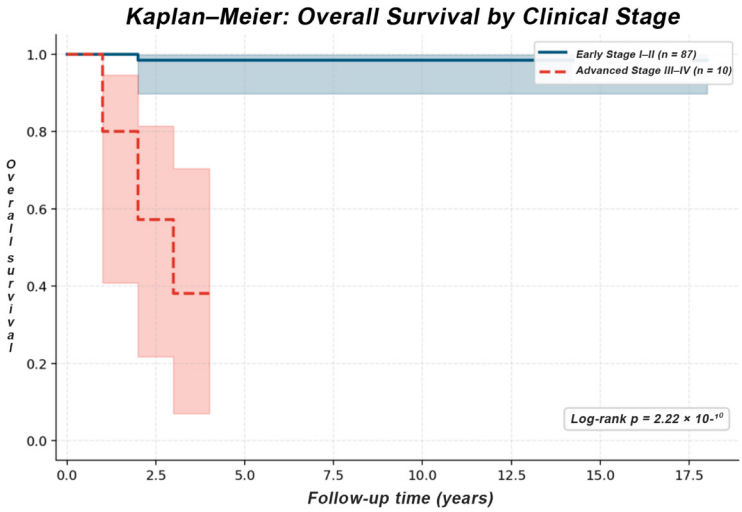

Figure 1. Kaplan–Meier curves of overall survival according to clinical stage. The advanced-stage group (III–IV; n = 10) showed significantly lower survival than the early-stage group (I–II; n = 87). Log-rank $p = 2.22 \times 10^{-10}$.

Survival by Triple-Negative Subtype

Figure 2 presents the Kaplan–Meier curves stratified by triple-negative subtype. The triple-negative group (n = 15) showed lower overall survival than the non-triple-negative group (n = 82), with a proportion of deaths of 26.7% vs. 2.4% (log-rank $p = 6.32 \times 10^{-4}$).

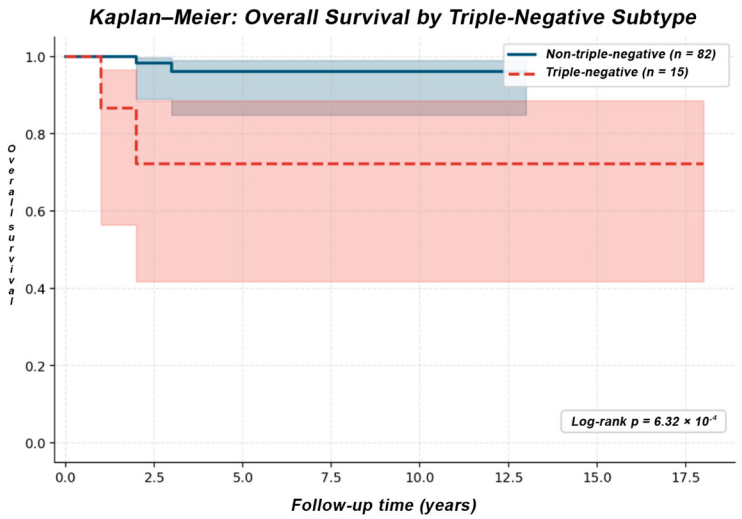

Figure 2. Kaplan–Meier curves of overall survival according to triple-negative subtype. The triple-negative group (n = 15) showed lower survival than the non-triple-negative group (n = 82). Log-rank $p = 6.32 \times 10^{-4}$.

Survival by Hormone Receptor Status

Figure 3 presents the Kaplan–Meier curves stratified by hormone receptor status. The HR-negative group (n = 22) showed poorer survival compared with the HR-positive group (n = 75), with a proportion of deaths of 22.7% vs. 1.3% (log-rank p = 6.27×10^{-4}).

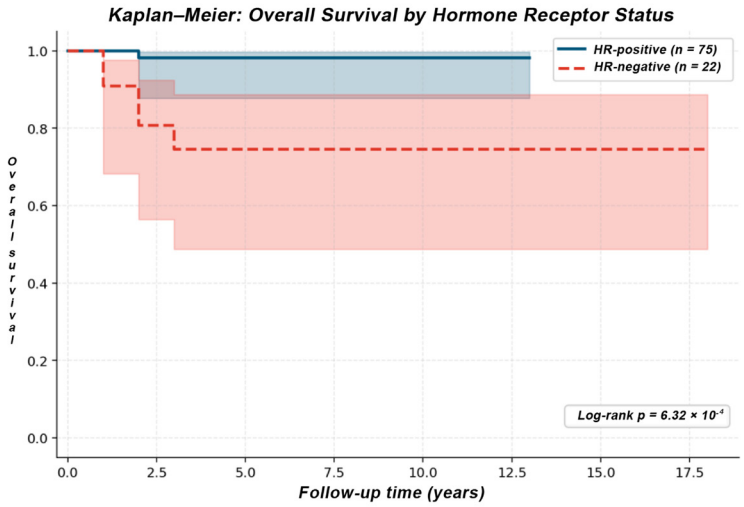

Figure 3. Kaplan–Meier curves of overall survival according to hormone receptor status. The HR-negative group (n = 22) showed lower survival than the HR-positive group (n = 75). Log-rank p = 6.27×10^{-4}.

Survival by HER2 Status

Figure 4 presents the Kaplan–Meier curves stratified by HER2 status. No significant difference in overall survival was observed between the HER2-positive (n = 19) and HER2-negative (n = 78) groups, with proportions of death of 5.3% and 6.4%, respectively (log-rank p = 0.983).

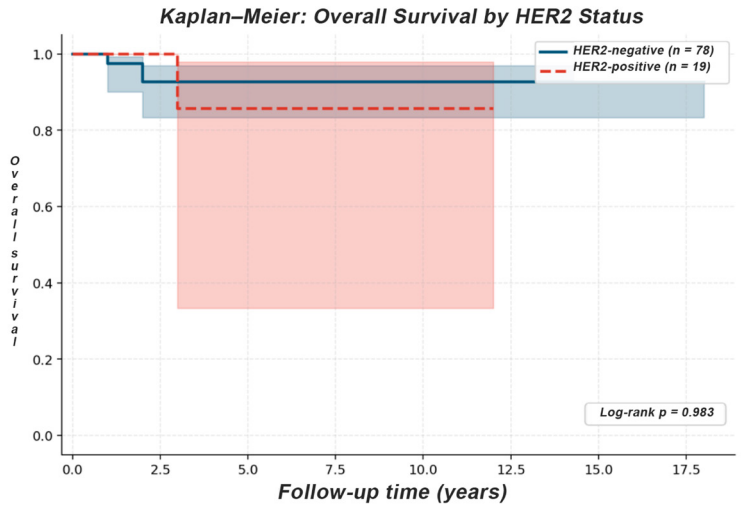

Figure 4. Kaplan–Meier curves of overall survival according to HER2 status. No significant difference was observed between the HER2-positive (n = 19) and HER2-negative (n = 78) groups. Log-rank p = 0.983.

DISCUSSION

The results of this analysis confirm well-established prognostic patterns, with a concentration of adverse events in the subgroups at higher clinical and molecular risk. The most pronounced difference was observed between the clinical stage groups: the proportion of deaths was 50% in the advanced-stage group vs. 1.1% in the early-stage group (log-rank $p = 2.22 \times 10^{-10}$). This gradient reflects both the greater tumor burden and the therapeutic limitations of locally advanced disease.[5]

The triple-negative subtype was associated with poorer survival (log-rank $p = 6.32 \times 10^{-4}$), with a proportion of deaths of 26.7% compared with 2.4% in the non-triple-negative group. This finding is consistent with the literature characterizing this subtype as having the most aggressive behavior among the molecular subtypes, given the absence of targeted therapies and the higher rate of cellular proliferation. The HR-negative group also showed poorer survival (log-rank $p = 6.27 \times 10^{-4}$), with 22.7% deaths vs. 1.3% in the HR-positive group — a result directly related to the overlap between HR negativity and biologically more aggressive subtypes, as well as the absence of benefit from adjuvant hormone therapy.[6]

The absence of a significant difference according to HER2 status ($p = 0.983$) is a finding that warrants contextualized interpretation. In private healthcare settings with access to anti-HER2 therapies — such as trastuzumab — the adverse prognostic impact of HER2 overexpression tends to be substantially attenuated by targeted treatment. The similar proportion of deaths between the HER2-positive (5.3%) and HER2-negative (6.4%) groups may therefore reflect the therapeutic benefit in this subgroup.[7]

The main limitations are the small number of events (6 deaths), which limits statistical power and the stability of survival estimates in the smaller subgroups, and the estimation of follow-up time based on the year of treatment — without precise dates of diagnosis and death — which introduces imprecision into the curves. These limitations are inherent to the retrospective nature and size of the cohort, and the results should be interpreted as exploratory.[8-10]

CONCLUSION

Advanced stage, the triple-negative subtype, and hormone receptor negativity were associated with poorer overall survival in this cohort of patients with invasive breast carcinoma, with statistically significant differences according to the log-rank test. HER2 status did not show a differential impact on survival, possibly due to the use of anti-HER2 therapies at the institution. These findings reinforce the importance of molecular stratification and early diagnosis for individualized prognostic assessment in breast oncology. Larger sample sizes and prospective data collection with precise dates are needed for more robust survival analyses.

REFERENCES

1. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457-81.

2. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep. 1966;50(3):163-70.

3. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.

4. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest. 2011;121(7):2750-67.

5. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers. 2019;5(1):66.

6. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-49.

7. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med. 2001;344(11):783-92.

8. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet. 2011;378(9793):771-84.

9. Gradishar WJ, Moran MS, Abraham J, Aft R, Agnese D, Allison KH, et al. NCCN Guidelines® Insights: Breast Cancer, Version 4.2023. J Natl Compr Canc Netw. 2023;21(6):594-608.

10. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2023. Disponível em: https://www.R-project.org/.

MAILING ADDRESS

JUAREZ ANTÔNIO DE SOUSA
Rua 13, 380 ap 5, Setor Oeste, Goiânia - GO, Brazil.
E-MAIL: Juarez_antonio@ufg.br

EDITORIAL AND REVIEW

Chief editors:

Waldemar Naves do Amaral - http://lattes.cnpq.br/4092560599116579 - https://orcid.org/0000-0002-0824-1138
Tárik Kassem Saidah - http://lattes.cnpq.br/7930409410650712- https://orcid.org/0000-0003-3267-9866

Authors:

JJuarez Antônio de Sousa - http://lattes.cnpq.br/4484429936026476 - https://orcid.org/0000-0001-5986-7926

Waldemar Naves do Amaral - http://lattes.cnpq.br/4092560599116579 - https://orcid.org/0000-0002-0824-1138

André Maroccolo de Sousa - http://lattes.cnpq.br/4379236388126901 - https://orcid.org/0009-0004-5613-1410

Débora Nascimento Dias Neves - http://lattes.cnpq.br/5830809552139767 - https://orcid.org/0000-0003-1280-5415

Laura de Sousa e Sousa - http://lattes.cnpq.br/5775215940326611- https://orcid.org/0009-0004-9019-8846

Luiza de Sousa e Sousa - http://lattes.cnpq.br/9882323304990808 - https://orcid.org/0009-0007-3855-8682

Isabella Rodrigues Ferreira - http://lattes.cnpq.br/4138419344172012 - https://orcid.org/0009-0001-9090-3814

Pâmela Christinny Fernandes Viêra - http://lattes.cnpq.br/4193211342761772 - https://orcid.org/0009-0002-1684-637X

Library Review: Izabella Goulart
Spell Check: Dario Alvares
Translation: Soledad Montalbetti Magri
Received: 31/05/26. Accepted: 14/07/26. Published in: 17/08/2026.