

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Instituto de Ciências Biomédicas
Programa de Pós-graduação em Imunologia e Parasitologia Aplicadas

IMPACTO DOS SUBTIPOS MOLECULARES E VARIÁVEIS CLÍNICO-PATOLÓGICAS
NOS DESFECHOS DO CÂNCER DE MAMA: UM ESTUDO ENTRE GRANDES BASES DE
DADOS

Gisele Cristiane da Silva Dias

Uberlândia-MG
2026

GISELE CRISTIANE DA SILVA DIAS

IMPACTO DOS SUBTIPOS MOLECULARES E VARIÁVEIS CLÍNICO-PATOLÓGICAS
NOS DESFECHOS DO CÂNCER DE MAMA: UM ESTUDO ENTRE GRANDES BASES DE
DADOS

Dissertação apresentada ao colegiado do Programa de Pós-graduação em Imunologia e Parasitologia Aplicadas da Universidade Federal de Uberlândia como requisito para obtenção do título de Mestre.

Gisele Cristiane da Silva Dias

Marcelo José Barbosa Silva

Uberlândia – MG
2026

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

D541
2026 Dias, Gisele Cristiane da Silva, 1984-
Impacto dos subtipos moleculares e variáveis clínico-patológicas
nos desfechos do câncer de mama [recurso eletrônico] : um
estudo entre grandes bases de dados / Gisele Cristiane da Silva
Dias. - 2026.

Orientador: Marcelo José Barbosa Silva.

Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Pós-graduação em Imunologia e Parasitologia Aplicadas.

Modo de acesso: Internet.

DOI <http://doi.org/10.14393/ufu.di.2026.408>

Inclui bibliografia.

Inclui ilustrações.

1. Imunologia. I. Silva, Marcelo José Barbosa, 1976-, (Orient.). II.
Universidade Federal de Uberlândia. Pós-graduação em Imunologia
e Parasitologia Aplicadas. III. Título.

CDU: 612.017

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

Nelson Marcos Ferreira - CRB6/3074



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Imunologia e Parasitologia Aplicadas				
Defesa de:	Dissertação de Mestrado Acadêmico, número 325, do PPGIPA				
Data:	Seis de julho de dois mil e vinte e seis	Hora de início:	14:25 h	Hora de encerramento:	17 h e 00 min
Matrícula do Discente:	12422IPA004				
Nome do Discente:	Gisele Cristiane da Silva Dias				
Título do Trabalho:	Impacto dos subtipos moleculares e variáveis clínico-patológicas nos desfechos do câncer de mama: um estudo entre grandes bases de dados				
Área de concentração:	Imunologia e Parasitologia Aplicadas				
Linha de pesquisa:	Imunobiologia dos processos alérgicos, autoimunes e tumorais				
Projeto de Pesquisa de vinculação:	Investigação do Perfil de Promotores Relacionados à Cancerização das Células Epiteliais de Mama				

Reuniu-se, em formato remoto pelo, pelo *Microsoft Teams*, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Imunologia e Parasitologia Aplicadas (ColPPGIPA), assim composta: Dr. Morun Bernardino Neto - FAMED/UFU; Dra. Lara Andrade Marques - Grupo Luta Pela Vida - GLPV e Dr. Marcelo José Barbosa Silva - ICBIM/UFU, presidente e orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Marcelo José Barbosa Silva, apresentou a Comissão Examinadora e o(a) candidato(a), agradeceu a presença do público, e concedeu à(ao) Discente a palavra para a exposição do seu trabalho. A duração da apresentação do(a) Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado(a).

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre(a).

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

3



Documento assinado eletronicamente por **Morun Bernardino Neto, Usuário Externo**, em 06/07/2026, às 17:03, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo José Barbosa Silva, Professor(a) do Magistério Superior**, em 06/07/2026, às 17:04, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Lara de Andrade Marques, Usuário Externo**, em 06/07/2026, às 17:07, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **7403351** e o código CRC **11510326**.

Referência: Processo nº 23117.039522/2026-37

SEI nº 7403351

AGRADECIMENTOS

À minha família, ao meu orientador e aos membros da banca examinadora, minha sincera gratidão pela contribuição à realização desse trabalho.

LISTA DE ABREVIATURAS E SIGLAS

AJCC	<i>American Joint Committee on Cancer</i>
BRCA	<i>Breast Cancer Genes</i>
CAAE	Certificado de Apresentação para Apreciação Ética
CAR-T	<i>Chimeric Antigen Receptor T cell</i>
CCR	<i>Canadian Cancer Registry</i>
CCCR	<i>Canadian Council of Cancer Registries</i>
CDC	<i>Centers for Disease Control and Prevention</i>
CID-O	Classificação Internacional de Doenças para Oncologia
CIF	Função de Incidência Cumulativa
CTLA-4	<i>Cytotoxic T-Lymphocyte Associated Protein 4</i>
CEP-UFU	Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Uberlândia
DAG	<i>Directed Acyclic Graph</i>
ECIS	<i>European Cancer Information System</i>
ENCR	<i>European Network of Cancer Registries</i>
EPV	Eventos por Variável
RE	Receptor de Estrogênio
ERBB2	Gene codificante de HER2
FOSP	Fundação Oncocentro de São Paulo
GVIF	<i>Generalized Variance Inflation Factor</i>
HER2	Receptor do fator de crescimento epidérmico humano tipo 2
IARC	<i>International Agency for Research on Cancer</i>
IBGE	Instituto Brasileiro de Geografia e Estatística

IDH	Índice de Desenvolvimento Humano
IHQ	Imuno-histoquímica
INCA	Instituto Nacional do Câncer
JRC	<i>Joint Research Centre</i>
Ki-67	Proteína nuclear expressa pelo gene MKI67
METABRIC	<i>Molecular Taxonomy of Breast Cancer International Consortium</i>
NCCR	<i>National Childhood Cancer Registry</i>
NCI	<i>National Cancer Institute</i>
NPCR	<i>National Program of Cancer Registries</i>
OMS	Organização Mundial da Saúde
PAM50	<i>Prediction Analysis of Microarray 50</i>
PD-1	<i>Programmed Cell Death Protein 1</i>
PD-L1	<i>Programmed Cell Death Ligand 1</i>
RA	Receptor de Andrógeno
RCBP	Registro de Câncer de Base Populacional
RHC	Registro Hospitalar de Câncer
RP	Receptor de Progesterona
SEER	<i>Surveillance, Epidemiology, and End Results</i>
StatCan	<i>Statistics Canada</i>
STROBE	<i>Strengthening the Reporting of Observational Studies in Epidemiology</i>
SUS	Sistema Único de Saúde
TNM	Sistema de Classificação de Tumores Malignos
UICC	<i>Union for International Cancer Control</i>

LISTA DE FIGURAS

1	Fluxo de seleção e padronização de dados	29
2	Comparação da disponibilidade de informações complementares nos registros analisados	29
3	Detalhamento dos registros incluídos na pesquisa nas fontes analisadas.	30
4	Avaliação preliminar do número de eventos para modelagem multivariada.	31
5	Fluxo de pré-processamento dos dados para avaliação da recorrência. . .	35
6	Incidência cumulativa de recorrência por subtipo intrínseco classificado pelo PAM50	43
7	Fluxo de preparação dos dados para análise comparativa entre registros.	48
8	Curvas ajustadas de sobrevida global e específica em 5 anos, segundo estadiamento tumoral	54
9	Fluxo análise dos registros de câncer de mama no RHC Nacional.	56
10	Perfil regional de utilização das modalidades de tratamento para o câncer de mama no Brasil (Período: 2000 a 2015)	59
11	Distribuição das distâncias percorridas por indivíduos que realizaram tratamento oncológico fora de seu município de residência. (Período: 2000 a 2015)	61
12	Mapeamento dos polos de tratamento e fluxos de deslocamento para terapias convencionais em pacientes com câncer de mama (Período: 2000 a 2015)	63
13	Mapeamento dos polos de tratamento e fluxos de deslocamento para imunoterapia em pacientes com câncer de mama (Período: 2000 a 2015)	64
B1	Grafo Acíclico Direcionado (DAG): modelo estrutural para avaliação da recorrência e riscos competitivos	101
B2	Resíduos do modelo de recorrência Fine–Gray (0–120 meses)	104
C1	Grafo Acíclico Direcionado (DAG): Modelo estrutural dos fatores clínicos associados ao óbito	106
C2	RHC Estadual: Resíduos do modelo de sobrevida global (Cox)	110
C3	RHC Estadual: Resíduos do modelo de sobrevida específica (Cox)	111
C4	RCBP Internacional: Resíduos do modelo de sobrevida global (Cox) . .	112
C5	RCBP Internacional: Resíduos do modelo de sobrevida específica (Cox)	113

LISTA DE TABELAS

1	Caracterização molecular pela imuno-histoquímica	33
2	Caracterização molecular por subtipos intrínsecos classificados pelo PAM50	34
3	Caracterização clínica e terapêutica da coorte molecular segundo os subtipos intrínsecos classificados pelo PAM50	37
4	Análise univariada Fine-Gray para o efeito total dos preditores sobre a recorrência	41
5	Análise multivariada Fine-Gray para o efeito ajustado dos preditores sobre a recorrência	42
6	Incidência cumulativa de recorrência do câncer de mama por subtipo intrínseco	44
7	Distribuição dos eventos e características do tempo de seguimento para diagnósticos realizados entre 2000 e 2015	46
8	Caracterização demográfica e clínica das coortes: RHC Estadual e RCBP Internacional (Período: 2000–2015)	47
9	Análise de sobrevida global em 5 anos (Cox): fatores associados ao óbito no RHC Estadual.	50
10	Análise de sobrevida global em 5 anos (Cox): Fatores associados ao óbito no RCBP Internacional	51
11	Análise de sobrevida específica por câncer em 5 anos (Cox causa-específico): Fatores associados ao óbito no RHC Estadual	52
12	Análise de sobrevida específica por câncer em 5 anos (Cox causa-específico): Fatores associados ao óbito no RCBP Internacional	53
13	Probabilidade de sobrevida em 5 anos segundo estadiamento tumoral . .	55
14	Caracterização demográfica e clínica do RHC Nacional (Período: 2000 a 2015)	57
15	Frequência de utilização das modalidades terapêuticas no RHC Nacional (Período: 2000 a 2015)	58
16	Distância percorrida (km) por indivíduos vinculados aos registros de câncer e tratados fora do município de residência (Período: 2000 a 2015)	60
B1	Diagnóstico de multicolinearidade dos preditores de recorrência	102
B2	Modelo multivariado de riscos proporcionais de Cox causa-específico para o período completo e janelas temporais	102

B3	Diagnóstico do pressuposto da proporcionalidade de riscos no modelo de recorrência	103
B4	Distribuição dos eventos por variáveis categóricas e períodos de seguimento	103
C1	Diagnóstico de multicolinearidade dos preditores de sobrevida	107
C2	Análise de Sensibilidade: Comparação de Hazard Ratios (IC 95%) para o RHC Estadual.	107
C3	Análise de Sensibilidade: Comparação de Hazard Ratios (IC 95%) para o RCBP Internacional.	108
C4	Diagnóstico do pressuposto da proporcionalidade de riscos nos modelos de sobrevida global	108
C5	Diagnóstico do pressuposto da proporcionalidade de riscos nos modelos de sobrevida específica	109
C6	Análise de sobrevida: contagem de eventos (óbitos) por tipo de registro .	109
D1	Frequência das modalidades terapêuticas segundo a região de tratamento no RHC Nacional	114
D2	Distribuição dos deslocamentos para tratamento oncológico fora da região de residência, por modalidade e segundo as regiões e faixas de distância (2000–2015)	115

SUMÁRIO

1	INTRODUÇÃO	14
1.1	A mama e o câncer	14
1.2	Classificação	15
1.3	Codificação CID-O	17
1.4	Incidência e fatores de risco	17
1.5	Diagnóstico e tratamento	19
1.6	Populações vulneráveis	21
1.7	Sobrevida, recorrência e riscos competitivos	22
1.8	O uso de informações contidas em bases de dados	22
1.9	Justificativa	24
2	OBJETIVOS	24
2.1	Geral	24
2.2	Específicos	25
3	METODOLOGIA	25
3.1	Desenho do estudo	25
3.2	Fontes de dados	26
3.3	Participantes e critérios de inclusão	27
3.4	Crterios de exclusão	27
3.5	Padronização	28
3.6	Tamanho da amostra	30
3.7	Análise Estatística	31
4	RESULTADOS	33
4.1	Caracterização dos perfis imuno-histoquímicos e genômicos	33
4.2	Caracterização clínica e terapêutica da coorte molecular segundo os subtipos intrínsecos	34
4.3	Fatores prognósticos associados ao risco de recorrência do câncer de mama	38
4.4	Análise da incidência cumulativa	42
4.5	Avaliação do tempo de seguimento para as análises de sobrevida em 5 anos	45
4.6	Caracterização das coortes: hospitalar e populacional	46
4.7	Análise comparativa dos fatores prognósticos entre registros de câncer .	48

4.8	Comparação das probabilidades de sobrevida global e específica em 5 anos	53
4.9	Caracterização demográfica, clínico-patológica e terapêutica dos Registros Hospitalares de Câncer do Brasil	55
4.9.1	Perfil demográfico e clínico	56
4.9.2	Distribuição dos tratamentos realizados	57
4.9.3	Distâncias percorridas para tratamento oncológico	59
5	Discussão	65
5.1	A robustez dos registros moleculares	65
5.2	A dinâmica temporal dos riscos em estudos oncológicos	69
5.3	Sobrevida global e específica no câncer de mama	70
5.4	Mapeamento geográfico e desigualdades na distribuição de serviços oncológicos no Brasil	72
5.5	Registros de câncer e desafios estruturais	74
6	Limitações do estudo	75
7	Conclusão	76
	APÊNDICES E ANEXOS	94
	Anexo A - Parecer ético	95
	Anexo B - STROBE checklist	98
	Apêndice A - Instrumento de Coleta de Dados	100
	Apêndice B - Diagnóstico de suposições e validação do modelo de recorrência do tumor	101
	Apêndice C - Diagnóstico de suposições e validação do modelo de sobrevida	106
	Apêndice D - Tratamento do câncer de mama no RHC Nacional	114
	Apêndice E - Declaração de Uso de Inteligência Artificial Generativa (IAG)	116

RESUMO

O câncer de mama representa um grande desafio de saúde global, e o uso de bases de dados públicas pode contribuir para uma compreensão mais abrangente dos padrões da doença. Considerando o potencial dessas fontes de dados para a realização de análises complementares, este trabalho teve como objetivo explorar diferentes bases de dados de acesso público para investigar características demográficas, tumorais e terapêuticas do câncer de mama. Realizou-se um estudo observacional e retrospectivo em registros hospitalares, populacionais e moleculares de quatro bases de dados: os Registros Hospitalares de Câncer (RHC de São Paulo e RHC Nacional); o Registro de Câncer de Base Populacional (RCBP) norte-americano SEER; e os dados genômicos e clínicos do METABRIC. Juntas, as quatro bases totalizaram 1.128.269 registros. Os diagnósticos ocorreram entre 2000 e 2015, exceto na base METABRIC. Empregou-se a modelagem de riscos competitivos de Fine-Gray para avaliar a recorrência tumoral em 120 meses, segmentada em janelas temporais (0 a 60 e 61 a 120 meses). Modelos de risco proporcional de Cox foram aplicados para estimar a sobrevida global e específica em 5 anos. A distância de deslocamento rodoviário necessária para tratamento no Brasil foi mensurada por meio dos pacotes estatísticos *geobr* e *osrm*. As análises foram realizadas no ambiente computacional RStudio, versão 2026.01.1+403. Os preditores independentes de risco médio de recorrência tumoral foram o tamanho do tumor, o número de linfonodos positivos, o grau histológico 3 e os subtipos moleculares intrínsecos definidos pela assinatura de expressão gênica PAM50. A análise da dinâmica temporal revelou que o comportamento do risco de recorrência tumoral não foi constante ao longo do tempo. Para a sobrevida em 5 anos, a idade e o estadiamento tumoral foram preditores independentes do risco médio de morte global e específica. O ano de diagnóstico mostrou associação inversa com o risco de morte específica, funcionando como marcador temporal de mudanças ocorridas ao longo do período estudado. Observou-se que, a cada ano mais recente de diagnóstico, o risco de morte específica diminuiu, em média, 2,81% no RHC de São Paulo e de 2,96% no RCBP. No RHC Nacional, constatou-se que 52,09% da amostra realizou tratamento fora do município de residência. As distâncias rodoviárias medianas calculadas por modalidade de tratamento variaram entre 76,6 km (hormonioterapia) e 114,0 km (imunoterapia). O mapeamento geográfico apontou desigualdades na distribuição dos serviços, demonstrando que os polos ofertantes de terapias estão densamente concentrados nas regiões Sudeste, Sul e Nordeste do Brasil. O estudo conclui que diferentes bases de dados fornecem informações complementares que ampliam o potencial analítico para a investigação dos desfechos do câncer de mama, possibilitando análises que não seriam contempladas por uma única fonte de dados. Esses achados reforçam o potencial de utilização conjunta de informações para apoiar pesquisadores e gestores de saúde no monitoramento dos padrões da doença e no planejamento de estratégias de atenção oncológica.

Palavras-chave: neoplasias da mama; tempo de recorrência; análise de sobrevida; terapêutica; registros hospitalares.

ABSTRACT

Breast cancer represents a major global health challenge, and the use of publicly available databases can contribute to a more comprehensive understanding of disease patterns. Given the potential of these data sources to support complementary analyses, this study aimed to explore different publicly available databases to investigate demographic, tumor, and therapeutic characteristics of breast cancer. An observational, retrospective study was conducted using hospital-based, population-based, and molecular records from four databases: the São Paulo Hospital-Based Cancer Registry (HBCR), the National HBCR, the Surveillance, Epidemiology, and End Results (SEER) Program, and the genomic and clinical data from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC). Together, these four databases comprised 1,128,269 breast cancer records. Diagnoses were made between 2000 and 2015, except for the METABRIC database. Fine-Gray competing risks modeling was employed to evaluate tumor recurrence over a 120-month follow-up period, stratified into two time intervals (0–60 and 61–120 months). Cox proportional hazards models were applied to estimate 5-year overall and cause-specific survival. Road travel distances to cancer treatment facilities in Brazil were estimated using the *geobr* and *osrm* statistical packages. All analyses were performed in the RStudio computing environment, version 2026.01.1+403. Independent predictors of the average risk of tumor recurrence included tumor size, the number of positive lymph nodes, histological grade 3, and intrinsic molecular subtypes defined by the PAM50 gene expression signature. Temporal dynamics analysis demonstrated that the risk of tumor recurrence was not constant over time. Regarding 5-year survival, age and tumor stage were independent predictors of the average risk of overall and cause-specific mortality. Year of diagnosis was inversely associated with the risk of cause-specific mortality, serving as a temporal marker of changes occurring throughout the study period. An average reduction of 2.81% in the São Paulo HBCR and 2.96% in the population-based registry was observed in the risk of cause-specific mortality for each more recent year of diagnosis. In the National HBCR, 52.09% of the sample received treatment outside their municipality of residence. Median road travel distances by treatment modality ranged from 76.6 km for hormone therapy to 114.0 km for immunotherapy. Geographic mapping revealed inequalities in the distribution of oncology services, demonstrating that treatment centers were highly concentrated in the Southeast, South, and Northeast regions of Brazil. The study concludes that different databases provide complementary information that expands the analytical capabilities for investigating breast cancer outcomes, enabling analyses that would not be possible using a single data source. These findings highlight the value of combining information from multiple data sources to support researchers and healthcare decision-makers in monitoring disease patterns and planning oncology care strategies.

Keywords: breast neoplasms; survival analysis; neoplasm recurrence; therapeutics; hospital-based cancer registries.

1 INTRODUÇÃO

1.1 A mama e o câncer

A glândula mamária está localizada bilateralmente na região anterior do tórax, sobre o músculo peitoral maior, estendendo-se em direção à axila ([HORTOBAGYI *et al.*, 2017](#)). Estruturalmente, é composta por dois compartimentos: o epitelial, constituído por lobos e ductos, organizados em unidades ducto-lobulares e revestidos por células epiteliais lumbais e mioepiteliais. O segundo compartimento, denominado estromal, é constituído por matriz extracelular, fibroblastos, adipócitos, células do sistema imune e vasos sanguíneos e linfáticos ([BISWAS *et al.*, 2022](#)).

Até a puberdade, o desenvolvimento das glândulas mamárias é semelhante entre indivíduos do sexo masculino e do feminino. Com o início da fase puberal, a ação hormonal do estrogênio, da progesterona e do hormônio do crescimento promove o desenvolvimento das mamas no sexo feminino, enquanto, no sexo masculino, a glândula mamária permanece como estrutura vestigial. Esse processo de desenvolvimento e diferenciação da glândula mamária é regulado por fatores hormonais e celulares e, quando alterado, pode contribuir para o surgimento de doenças como o câncer de mama ([JAVED; LTEIF, 2013](#); [MCSHERRY *et al.*, 2007](#)).

O câncer pode ser descrito como o resultado de um acúmulo gradual de alterações genéticas e epigenéticas ao longo do tempo, que permitem à célula tumoral escapar de mecanismos regulatórios de proliferação e sobrevivência ([TAKESHIMA; USHIJIMA, 2019](#); [VERONA *et al.*, 2022](#)). Em condições normais, as células estromais do tecido mamário podem restringir o crescimento tumoral por meio da ativação de genes supressores ([CHO, 2022](#); [JANIC; ABAD; AMELIO, 2025](#)). Contudo, a falha nesse processo e, conseqüentemente, a interrupção da homeostase tecidual podem ser descritas em três etapas:

1. **Iniciação:** ocorre após a exposição a agentes carcinogênicos (químicos, físicos ou biológicos), durante a qual a célula pode sofrer alterações genéticas irreversíveis. Se os mecanismos de reparo ou de apoptose não corrigirem a mutação, as alterações podem se perpetuar nas divisões celulares subsequentes ([KOYA; IBRAHIM, 2025](#)).
2. **Promoção:** nesta etapa, as células iniciadas são estimuladas a proliferar por agentes promotores, como hormônios e mediadores inflamatórios, que aceleram a progressão dessas células, levando à formação de um tumor detectável por células do sistema imune ([SEVER; BRUGGE, 2015](#)).
3. **Progressão:** as mutações se acumulam e há uma interação complexa entre células tumorais, células estromais, células do sistema imune e matriz extracelular, formando

um microambiente que sustenta o crescimento do tumor, modulando o comportamento, a progressão da doença e a resposta ao tratamento (DE VISSER; JOYCE, 2023).

À medida que o câncer progride, surgem modificações morfológicas e histológicas que refletem a complexidade biológica do câncer de mama, com diferentes padrões de diferenciação celular e comportamentos clínicos.

1.2 Classificação

Dada a heterogeneidade e a complexidade da doença, o câncer de mama pode ser classificado em diferentes tipos com base em múltiplos critérios. A seguir, serão apresentadas as quatro principais categorias de classificação:

1. **Origem:** a patogênese do câncer de mama inicia-se com a proliferação de células epiteliais anormais, originadas predominantemente nos ductos ou lóbulos, resultando em lesões pré-invasivas denominadas carcinomas *in situ*. A progressão para o carcinoma invasivo ocorre quando essas células tumorais rompem a membrana basal e infiltram o estroma adjacente (NOLAN; LINDEMAN; VISVADER, 2023). Embora a grande maioria dos casos se origine do epitélio ductolobular, existem formas raras de neoplasias mamárias (SONG; SINGH, 2024). Os tumores epiteliais raros descritos na literatura incluem os carcinomas medulares, mucinosos, tubulares, papilíferos, micropapilíferos, apócrinos, cribriformes, adenoescamosos, adenocísticos, neuroendócrinos, secretores, metaplásicos, lipídicos, oncocíticos e sebáceos, além da doença de Paget da mama, que apresentam características morfológicas e comportamentos biológicos distintos (BARRIENTOS-TORO; DING; RASO, 2024; MARKARIAN; HOLMES, 2022; SONG; SINGH, 2024). Além dos carcinomas epiteliais, há também neoplasias mamárias raras de origem não epitelial, derivadas do tecido conjuntivo ou vascular, como os tumores filóides, angiosarcomas e linfomas primários da mama (SONG; SINGH, 2024).
2. **Classificação molecular e perfis imuno-histoquímicos:** os perfis imuno-histoquímicos são obtidos por meio da Imuno-histoquímica (IHQ), técnica que avalia a expressão de biomarcadores proteicos como Receptor de Estrogênio (RE), Receptor de Progesterona (RP), Receptor do fator de crescimento epidérmico humano tipo 2 (HER2) e Proteína nuclear expressa pelo gene MKI67 (Ki-67) (ORRANTIA-BORUNDA *et al.*, 2022). Com base na combinação desses marcadores, os tumores são classificados em subtipos: Luminal A, Luminal B, HER2 positivo (Não Luminal) e Triplo Negativo (COATES *et al.*, 2015; GOLDBHIRSCH *et al.*, 2011). Paralelamente, avanços na transcriptômica possibilitaram o desenvolvimento de assinaturas de expressão gênica, como o *Prediction Analysis*

of Microarray 50 (PAM50), que classifica os tumores em quatro subtipos intrínsecos: Luminal A, Luminal B, HER2 enriquecido e *basal-like* (DE HAAS *et al.*, 2023; NOLAN; LINDEMAN; VISVADER, 2023; POMMIER *et al.*, 2020). Outros subtipos como *normal-like* e *claudin-low*¹ também foram descritos com base em perfis de expressão gênica, ampliando a compreensão da heterogeneidade biológica do câncer de mama (FOUGNER *et al.*, 2020; PARKER *et al.*, 2009; POMMIER *et al.*, 2020; VOUTSADAKIS, 2023).

3. **Grau histológico:** expressa o nível de diferenciação celular e reflete o grau de agressividade biológica. Ele é determinado por sistemas de classificação como o de *Bloom-Richardson* e sua versão modificada, o *Nottingham*, atualmente o mais utilizado. Ambos classificam os tumores com base em três critérios: formação de túbulos, pleomorfismo nuclear e índice mitótico (RAKHA; TOSS; QUINN, 2022). Cada critério recebe uma pontuação de 1 a 3, e a soma total define o grau histológico², classificando os tumores em: grau I, que indica tumor bem diferenciado; grau II, de diferenciação moderada; e grau III, pouco diferenciado (HORTOBAGYI *et al.*, 2017). Essa classificação é fundamental para prever o comportamento do tumor e orientar decisões terapêuticas (JAROENSRI *et al.*, 2022).
4. **Estadiamento:** baseado principalmente no Sistema de Classificação de Tumores Malignos (TNM), desenvolvido pela *Union for International Cancer Control* (UICC) e adotado pelo *American Joint Committee on Cancer* (AJCC) (RAMI-PORTA, 2024). O sistema é composto por três elementos principais (HORTOBAGYI *et al.*, 2017):
 - **T** se refere ao tumor primário e descreve o tamanho e a extensão da lesão mamária, onde T0 indica ausência de tumor detectável, enquanto T1 a T4 correspondem a estágios progressivos de crescimento e invasão;
 - **N** avalia os linfonodos regionais, determinando o número e a localização, onde N0 indica ausência de disseminação linfonodal, ao passo que N1 a N3 representam níveis progressivos de envolvimento dos linfonodos regionais;
 - **M** avalia a ocorrência de metástases à distância, em que M0 indica ausência de metástases e M1 indica disseminação para outros órgãos.

A combinação desses parâmetros permite a classificação clínica da extensão anatômica em estágios de I a IV, refletindo o avanço da doença (RAMI-PORTA, 2024).

¹Os tipos de tumores *claudin-low* (baixa expressão de claudina) representam um grupo adicional dentro dos carcinomas mamários, caracterizado por baixa expressão de genes relacionados às junções celulares como claudinas 3, 4 e 7, oclusina e E-caderina.

²O grau histológico é considerado grau I quando a soma dos critérios resulta em 3, 4 ou 5 pontos; grau II a soma resulta em 6 ou 7 pontos e grau III em 8 ou 9 pontos.

1.3 Codificação CID-O

A Classificação Internacional de Doenças para Oncologia (CID-O), atualmente em sua 4ª edição, é um sistema de codificação de tumores desenvolvido pela Organização Mundial da Saúde (OMS) e utilizado internacionalmente para a padronização dos registros de câncer (ZNAOR *et al.*, 2026). Esse sistema codifica tumores em duas categorias principais (ORGANIZATION, 2013):

- **Topografia:** indica a localização anatômica do tumor. Os códigos alfanuméricos são agrupados de C00.0 a C80.9.
- **Morfologia:** descreve o tipo histológico do tumor. Os códigos numéricos são agrupados de 8000/0 a 9989/3, sendo que o último dígito após a barra (/) indica o comportamento biológico, diferenciando os tumores em (/0) benignos, (/1) comportamento incerto, (/2) *in situ* e (/3) maligno considerado ou presumido como primário. Há ainda os códigos (/6) e (/9) categorizados como malignos, contudo, esses não são utilizados nos registros de câncer.

1.4 Incidência e fatores de risco

As mulheres são predominantemente afetadas e, embora a doença também possa acometer homens, sua ocorrência neste grupo é significativamente menor (BHARDWAJ *et al.*, 2024; BRAY *et al.*, 2024). Em 2026, estimou-se a ocorrência de 321.910 novos casos de câncer de mama em mulheres e 2.670 novos casos em homens, nos Estados Unidos (SIEGEL *et al.*, 2026). Globalmente, a incidência padronizada por idade foi estimada em 46,40 casos por 100.000 mulheres e 0,94 casos por 100.000 homens, segundo estimativas do *Burden of Disease Study 2021* (LI *et al.*, 2025; WU; XIA; LIN, 2024). No Brasil, para o triênio 2026-2028, estima-se 78.610 casos novos de câncer de mama feminino por ano, correspondendo a uma taxa ajustada de incidência de 42,50 casos por 100.000 mulheres, sendo as maiores taxas observadas na região Sudeste (51,72 por 100.000 mulheres) e Sul (47,65 por 100.000 mulheres) (MARTINS *et al.*, 2026).

O câncer de mama feminino decorre da interação entre fatores genéticos, hormonais e não genéticos. A predisposição genética é um dos elementos relevantes, sobretudo devido às mutações em genes como *Breast Cancer Genes* (BRCA) 1 e 2 (XIONG *et al.*, 2025). Esses genes codificam proteínas importantes, relacionadas ao reparo do DNA, e desempenham um papel importante na supressão do tumor. Mutações nesses genes podem ser herdadas, através das células germinativas, aumentando o risco de certos tipos de câncer (ARUN *et al.*, 2024).

Entre os fatores não genéticos, a literatura destaca a idade avançada, a história reprodutiva (menarca precoce, menopausa tardia, nuliparidade e primeira gestação em idade tardia), bem como o uso de terapia hormonal na pós-menopausa, todos relacionados a uma maior exposição cumulativa ao estrogênio (YIALLOUROU *et al.*, 2024; OBEAGU; OBEAGU, 2024). Ao passo que a gravidez precoce e a amamentação são relatadas como fatores protetores ou atenuadores (XIONG *et al.*, 2025). Além disso, o estilo de vida desempenha um papel significativo: obesidade, sedentarismo, consumo excessivo de álcool, tabagismo e dietas pouco saudáveis estão associados a um maior risco de desenvolvimento da doença (HOXHA *et al.*, 2024).

A maior ocorrência entre mulheres pode ser explicada, em parte, pela biologia do tecido mamário, mais desenvolvido e mais sensível à exposição hormonal cumulativa ao longo da vida reprodutiva (NOLAN; LINDEMAN; VISVADER, 2023). Os hormônios sexuais têm grande importância para a homeostase do epitélio mamário, principalmente no desenvolvimento da glândula mamária na puberdade e no período pós-natal (BRISKEN; N; O'MALLEY, 2010; MCSHERRY *et al.*, 2007). Contudo, os mesmos hormônios que regulam a normalidade do tecido mamário também afetam o curso da doença (KIM; MUNSTER, 2025).

O estrogênio desempenha um papel central no crescimento e na diferenciação dos ductos mamários, enquanto a progesterona estimula a maturação dos lóbulos e alvéolos mamários. A prolactina atua na diferenciação terminal e no preparo para a lactação, e a ocitocina é essencial para a ejeção do leite, regulando a função secretória da glândula (BRISKEN; N; O'MALLEY, 2010; MCSHERRY *et al.*, 2007). Mais recentemente, os andrógenos têm despertado interesse por sua ação moduladora sobre a proliferação epitelial e pela interação com receptores hormonais (RATHS *et al.*, 2023). Revisões recentes destacam o papel do estrogênio na carcinogênese mamária, ao modular a transcrição, a inflamação e os danos ao DNA ao longo da vida, conectando exposições endógenas e exógenas ao risco individual de câncer de mama feminino (KIM; MUNSTER, 2025; MOHANTY *et al.*, 2024; NOLAN; LINDEMAN; VISVADER, 2023).

Em comparação com o câncer de mama feminino, o câncer de mama masculino é considerado incomum. Essa diferença marcante pode ser associada principalmente a atributos biológicos relacionados à regulação hormonal. Embora rudimentar, o tecido mamário de indivíduos do sexo masculino expressa receptores hormonais, da mesma forma que indivíduos do sexo feminino, incluindo RE, RP, Receptor de Andrógeno (RA) e HER2 (CARDOSO *et al.*, 2018). No entanto, o estímulo estrogênico, que promove o crescimento ductal, é normalmente regulado pela ação dos andrógenos e, como consequência, a mama masculina apresenta-se involuída, composta predominantemente por tecido adiposo subcutâneo, com remanescentes de tecido ductal pela ação do estrogênio (CHARLOT *et al.*, 2013; SHI; XIN, 2024). Alterações nessa dinâmica regulatória podem levar à proliferação anômala do tecido mamário, aumentando a suscetibilidade ao desenvolvimento de câncer de mama masculino (BHARDWAJ *et al.*, 2024).

A idade avançada é um importante fator associado à ocorrência do câncer de mama masculino, com maior incidência observada após os 60 anos (LI *et al.*, 2025). Fatores genéticos e

hereditários, como mutações nos genes [BRCA1/BRCA2](#), bem como condições clínicas associadas ao hiperestrogenismo (incluindo síndrome de Klinefelter, hipogonadismo e obesidade), também estão implicados na patogênese da doença ([CHIDAMBARAM et al., 2024](#)).

1.5 Diagnóstico e tratamento

O diagnóstico do câncer de mama deve seguir uma abordagem integrada, que combina avaliação clínica detalhada, exames de imagem e análises histopatológicas e imuno-histoquímicas. A investigação clínica inclui a coleta da história clínica do indivíduo e a realização de um exame físico minucioso. Em seguida, são solicitados exames de imagem, conforme a faixa etária e as características clínicas. A confirmação diagnóstica é obtida por meio de biópsia por agulha grossa guiada por imagem, permitindo a análise do tecido tumoral. Nessa etapa, são avaliados os biomarcadores [RE](#), [RP](#), [HER2](#) e [Ki-67](#), parâmetros fundamentais para o planejamento terapêutico e a previsão da resposta ao tratamento ([APHALE et al., 2023](#); [BRASIL, 2024](#)).

As modalidades terapêuticas no câncer de mama incluem abordagens neoadjuvantes, administradas antes do tratamento principal, e adjuvantes, aplicadas após o tratamento principal, sendo ambas escolhidas conforme o estadiamento do tumor, o perfil biológico, além das condições clínicas e das expectativas do indivíduo que receberá o tratamento ([BRASIL, 2024](#); [KERR et al., 2022](#)). Os tratamentos também podem ser classificados como locorregionais, quando atuam diretamente sobre o tumor e áreas adjacentes, como cirurgia e radioterapia, ou sistêmicos, quando visam combater células tumorais que possam ter se disseminado, como ocorre na quimioterapia, hormonioterapia, imunoterapia e terapias-alvo ([CHEN et al., 2024](#)).

Em muitos casos, a cirurgia constitui a primeira etapa do tratamento do câncer de mama, frequentemente associada a terapias adjuvantes ou neoadjuvantes. A abordagem cirúrgica no câncer de mama pode ter diferentes finalidades, sendo dividida em procedimentos terapêuticos e diagnósticos. No âmbito do diagnóstico, destacam-se a biópsia do linfonodo sentinela, que permite identificar a presença de metástases nos linfonodos axilares, e a dissecação axilar, indicada em casos selecionados e em casos de comprometimento linfonodal mais extenso ([WANG; WU, 2023](#)). Embora esses procedimentos tenham sido objeto de debate na literatura recente, sobretudo em razão dos riscos associados e da evolução das técnicas de imagem e dos marcadores prognósticos ([GENTILINI et al., 2023](#); [PARK et al., 2025](#); [TINTERRI et al., 2023](#)), ambas as abordagens continuam sendo amplamente utilizadas na prática clínica, contribuindo de forma relevante para o estadiamento tumoral. Já os procedimentos terapêuticos visam à remoção do tumor, como ocorre na cirurgia conservadora e na mastectomia. Além disso, outras intervenções cirúrgicas podem integrar o plano terapêutico, como a reconstrução mamária e a cirurgia paliativa ([LUTENCO et al., 2024](#)).

A quimioterapia é uma das principais modalidades terapêuticas no tratamento do câncer de mama e pode ser administrada em diferentes momentos ao longo do tratamento. A quimioterapia neoadjuvante é especialmente recomendada antes da cirurgia para pacientes com tumores grandes, linfonodos axilares múltiplos comprometidos e subtipos agressivos, como o câncer de mama triplo negativo e [HER2-positivo](#). Já a quimioterapia adjuvante é administrada após a cirurgia, com o intuito de reduzir o risco de recorrência ([WANG; WU, 2023](#)).

A radioterapia é uma modalidade terapêutica locorregional amplamente empregada no tratamento do câncer de mama em estágio inicial. Sua indicação é frequente após a cirurgia conservadora, embora também possa ser aplicada no contexto pós-mastectomia ([BRASIL, 2024](#); [WHELAN et al., 2023](#)). O mecanismo de ação da radioterapia baseia-se na emissão de radiação ionizante capaz de destruir células tumorais ou inibir sua multiplicação. Essa modalidade terapêutica pode ser classificada em: radioterapia de feixe externo, na qual a radiação é emitida por uma fonte localizada fora do corpo e direcionada ao tumor; e radioterapia de feixe interno (braquiterapia), que envolve a inserção de fontes radioativas próximas ou diretamente no tumor ([NATIONAL INSTITUTES OF HEALTH, 2023](#); [WANG; WU, 2023](#)).

A hormonioterapia, também denominada terapia endócrina, é considerada o pilar do tratamento do câncer de mama inicial com receptores hormonais positivos. Seu objetivo é bloquear o estímulo estrogênico responsável pelo crescimento tumoral, podendo ser aplicado em diferentes contextos e com distintas estratégias conforme o status menopausal e o risco de recorrência ([CUCCINIELLO et al., 2022](#)). Atualmente, testes genômicos como *Oncotype DX*, *PAM50* (Prosigna) e *MammaPrint* têm sido utilizados para avaliar a recorrência do câncer de mama e auxiliar na decisão quanto à necessidade de quimioterapia adjuvante, permitindo identificar quais pacientes podem ser tratados apenas com terapia endócrina ([ANDRE et al., 2022](#)).

Diferentemente das terapias convencionais, a imunoterapia é uma abordagem terapêutica inovadora que visa restaurar ou potencializar a resposta imune antitumoral, permitindo que o próprio sistema imunológico reconheça e elimine as células neoplásicas. Entre as principais modalidades, destacam-se os inibidores de *checkpoints* imunológicos, que bloqueiam moléculas reguladoras como *Programmed Cell Death Protein 1* ([PD-1](#)), *Programmed Cell Death Ligand 1* ([PD-L1](#)) e *Cytotoxic T-Lymphocyte Associated Protein 4* ([CTLA-4](#)), promovendo a reativação da função citotóxica dos linfócitos T e restabelecendo o controle imunológico sobre o tumor ([KCIUK et al., 2023](#)). Outra estratégia é a transferência adotiva de células T, incluindo as terapias com *Chimeric Antigen Receptor T cell* ([CAR-T](#)), nas quais linfócitos são geneticamente modificados e passam a expressar receptores quiméricos que reconhecem antígenos tumorais específicos e exercem ação antitumoral direcionada ([LEE; LEE, 2024](#)).

As vacinas terapêuticas contra o câncer têm ganhado destaque por estimularem o sistema imunológico a reconhecer e eliminar células malignas, com base em antígenos tumorais específicos expressos no tecido tumoral ([FAN et al., 2023](#)). Paralelamente, as terapias com

citocinas representam uma abordagem imunoterapêutica que visa intensificar a resposta imune antitumoral, por meio da ativação e proliferação de células efectoras, envolvidas na destruição das células neoplásicas (QIU *et al.*, 2021). Outra abordagem relevante é a viroterapia oncolítica, que utiliza vírus geneticamente modificados capazes de infectar e destruir seletivamente células tumorais, intensificando a resposta imunológica local (KCIUK *et al.*, 2023).

As terapias anti-HER2, também denominadas terapias-alvo, constituem uma classe de medicamentos desenvolvidos para o tratamento do câncer de mama HER2 positivo, abrangendo desde os estágios iniciais até a doença avançada (STANOWICKA-GRADA; SENKUS, 2023). No câncer de mama HER2 positivo, observa-se amplificação do Gene codificante de HER2 (ERBB2), responsável pela superexpressão da proteína HER2, um receptor transmembrana envolvido nos processos de crescimento e proliferação celulares (ROJAS-SALAZAR *et al.*, 2024). As terapias anti-HER2 foram desenvolvidas para bloquear a ativação anômala desses receptores, interrompendo as vias de sinalização que promovem o crescimento tumoral e, consequentemente, reduzindo a progressão da doença. Estudos recentes demonstraram que parte dos tumores anteriormente classificados como HER2 negativos apresenta níveis baixos, porém detectáveis, de expressão proteica, ampliando o espectro de pacientes potencialmente beneficiados por essas terapias direcionadas (ROY *et al.*, 2023; ZHU *et al.*, 2024).

1.6 Populações vulneráveis

Embora o câncer de mama tenha origem em alterações celulares específicas, sua manifestação e evolução podem ser significativamente influenciadas por fatores socioeconômicos e pelas desigualdades no acesso aos serviços de saúde, o que contribui para ampliar a vulnerabilidade em determinados grupos populacionais (SLADE *et al.*, 2024). Estudos realizados nos Estados Unidos e no Brasil indicam que mulheres negras são diagnosticadas com maior frequência em estágios mais avançados da doença (BRAWLEY; LANSEY, 2023; MARCELINO *et al.*, 2023). Essas desigualdades ainda se agravam devido a barreiras geográficas de acesso, com indivíduos residentes em regiões mais remotas percorrendo grandes distâncias para alcançar centros de tratamento oncológico (ADAMS *et al.*, 2025; FONSECA *et al.*, 2022).

Embora os contextos populacionais apresentem suas particularidades, fatores estruturais, socioeconômicos e educacionais ampliam as desigualdades no controle do câncer de mama. Nos Estados Unidos, os custos associados ao acesso a serviços especializados de tratamento oncológico têm sido apontados na literatura como uma barreira à logística do tratamento (PLANEY *et al.*, 2024). No Brasil, o impacto da segregação de renda tem se refletido na mortalidade por câncer de mama entre as mulheres (GUIMARÃES *et al.*, 2024), enquanto, na China, a associação entre o nível de escolarização das mulheres e o câncer de mama tem sido alvo de investigação (JIANG *et al.*, 2023). O próprio fato de a maioria dos estudos sobre o câncer de mama estar

direcionada à população feminina evidencia a menor visibilidade da população masculina, cujas particularidades clínicas e epidemiológicas ainda são pouco exploradas, o que pode contribuir para lacunas no conhecimento científico e nas práticas assistenciais.

1.7 Sobrevida, recorrência e riscos competitivos

A sobrevida e a recorrência tumoral constituem alguns dos principais desfechos empregados na avaliação prognóstica do câncer de mama, permitindo estimar a evolução da doença e subsidiar decisões terapêuticas. Tradicionalmente, esses desfechos são avaliados por métodos de análise de sobrevida, como o estimador de Kaplan-Meier e o modelo de riscos proporcionais de Cox ([STALPERS; KAPLAN, 2018](#)).

Entretanto, em muitos estudos oncológicos, especialmente aqueles que incluem indivíduos idosos ou com múltiplas comorbidades, podem ocorrer eventos que impedem a observação do desfecho de interesse, como o óbito por outras causas não relacionadas ao câncer. Esses eventos são denominados riscos competitivos e devem ser considerados nas análises de recorrência tumoral, pois influenciam a probabilidade de ocorrência desse desfecho ao longo do acompanhamento ([SALEH *et al.*, 2020](#)).

Nesse contexto, Fine e Gray (1999) propuseram um modelo de regressão para análise de riscos competitivos que permite estimar a incidência cumulativa do evento de interesse na presença de eventos competidores. Desde então, essa abordagem tem sido utilizada em estudos epidemiológicos e clínicos, onde o envelhecimento populacional torna os eventos competidores cada vez mais relevantes na avaliação prognóstica ([FINE; GRAY, 1999](#); [SØRENSEN; HORVÁTH-PUHÓ; PEACOCK, 2026](#)).

1.8 O uso de informações contidas em bases de dados

A aplicabilidade do uso de bases de dados na oncologia inclui, além da análise em grande escala, o monitoramento da incidência e da mortalidade, a avaliação de fatores de prognóstico e de sobrevida, a identificação de disparidades nos cuidados, o desenvolvimento de modelos preditivos de risco, entre outras aplicações ([BRAWLEY; LANSEY, 2023](#); [CASWELL-JIN *et al.*, 2024](#); [GUPTA *et al.*, 2024](#); [KUMAR *et al.*, 2021](#)).

As bases de dados são alimentadas por diversas instituições de saúde e, no que se refere à organização dos dados oncológicos, distinguem-se dois modelos de registro: os populacionais e os hospitalares. O Registro Hospitalar de Câncer ([RHC](#)) tem por objetivo coletar todos os casos de câncer diagnosticados e/ou tratados em uma instituição de saúde, permitindo monitorar

e avaliar a assistência prestada, o tratamento e a evolução clínica ([NATIONAL INSTITUTES OF HEALTH, n.d.\[b\]](#)). No Brasil, os hospitais especializados no atendimento oncológico são obrigados a manter o [RHC](#), conforme determina a Portaria nº 140, de 27 de fevereiro de 2014 ([BRASIL, 2014](#)). Já o Registro de Câncer de Base Populacional ([RCBP](#)) coleta dados de todas as instituições de saúde de uma área geográfica definida, possibilitando estimar a incidência real de câncer em determinada população ([PENA et al., 2025](#); [SILVA; OLIVEIRA; CURADO, 2024](#)).

Vinculada à [OMS](#), a *International Agency for Research on Cancer* ([IARC](#)) atua como referência global para os [RCBPs](#) e coordena iniciativas voltadas ao aprimoramento da coleta de dados de alta qualidade sobre o câncer ([PIÑEROS et al., 2017](#)). Estabelecido em 2014, o centro regional da [IARC](#) para a América Latina é composto por um centro de coordenação central e centros colaboradores, estrategicamente distribuídos na região. No Brasil, o Instituto Nacional do Câncer ([INCA](#)) atua como centro colaborador da [IARC](#), integrando os dados coletados sobre o câncer no território nacional ([INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. IARC, 2023](#)). A implantação dos [RCBPs](#) no Brasil teve início na década de 1960 e, nas décadas seguintes, o modelo foi gradualmente expandido para diversas capitais e centros urbanos, totalizando atualmente cerca de 30 [RCBPs](#) distribuídos pelo país ([NETO, 2021](#); [PIÑEROS et al., 2021](#)).

Nos Estados Unidos, os registros de câncer são integrados pelo *National Program of Cancer Registries* ([NPCR](#)), mantido pelo *Centers for Disease Control and Prevention* ([CDC](#)) e pelo *Surveillance, Epidemiology, and End Results* ([SEER](#)), mantido pelo *National Cancer Institute* ([NCI](#)) ([DUGGAN et al., 2016](#); [FRIEDMAN; NEGOITA, 2024](#)).

O fluxo de informações nos registros de câncer americano segue um padrão estruturado que assegura a qualidade e a abrangência dos dados. O processo inicia-se no momento em que o paciente recebe o diagnóstico, e as informações clínicas e demográficas são coletadas pelo registrador de câncer. Essas informações são posteriormente enviadas ao registro central estadual. Nesse nível, ocorre a consolidação de dados provenientes de diferentes hospitais e laboratórios do estado, com revisão e verificação para garantir a consistência e a completude. Anualmente, os registros centrais estaduais remetem os dados ao [NPCR](#), onde são integrados ao programa [SEER](#) e ao Sistema Nacional de Estatísticas Vitais do [CDC](#), constituindo uma base robusta que subsidia estatísticas, políticas públicas e estratégias de controle do câncer em nível nacional ([CENTERS FOR DISEASE CONTROL AND PREVENTION, 2024](#); [NATIONAL INSTITUTES OF HEALTH, n.d.\[b\]](#)).

No Canadá, os registros de câncer seguem um modelo populacional centralizado através do *Canadian Council of Cancer Registries* ([CCCR](#)), mantido pela *Statistics Canada* ([StatCan](#)). Esse sistema é alimentado localmente pelas províncias e territórios, que enviam ao *Canadian Cancer Registry* ([CCR](#)) os dados clínicos e demográficos dos indivíduos diagnosticados com câncer. A governança do [CCR](#) é realizada pelo [CCCR](#) e pelo Centro de Dados de Saúde da População da [StatCan](#) ([GOVERNMENT OF CANADA, 2025](#)).

Na Europa, o fluxo de dados sobre câncer inicia-se nos [RCBPs](#). A centralização dos dados é realizada por meio do *European Cancer Information System* ([ECIS](#)), que analisa e divulga os indicadores epidemiológicos em toda a União Europeia. A coordenação da validação e da consistência metodológica dos registros é conduzida pelo *Joint Research Centre* ([JRC](#)) e pela *European Network of Cancer Registries* ([ENCR](#)), que definem os protocolos técnicos, capacitam os profissionais de registros e atualizam as recomendações para padronização dos dados ([GIUSTI et al., 2024](#)).

Além dos registros populacionais e hospitalares, destacam-se as bases de dados genômicos, que têm ampliado significativamente o entendimento da biologia tumoral e de suas implicações clínicas. Um exemplo é o *cBioPortal for Cancer Genomics*, um recurso *online* de acesso aberto que permite visualizar e analisar dados genômicos relacionados ao câncer. O portal também oferece suporte à exploração de dados clínicos, à análise de sobrevida e ao *download* seletivo de informações, possibilitando ao usuário escolher entre os diversos estudos disponíveis na plataforma ([GAO et al., 2013](#)).

1.9 Justificativa

Diante do exposto, a utilização de informações de saúde provenientes de bases de dados contendo registros de câncer justifica-se por possibilitar análises abrangentes e sólidas dos fatores que influenciam os desfechos oncológicos. Essas análises podem contribuir para o entendimento da evolução da doença, oferecendo subsídios para aprimorar as estratégias de manejo clínico. Por fim, ao incorporar a análise da distribuição espacial dos serviços de saúde e do deslocamento para o tratamento, estudos como este podem orientar a formulação de estratégias de cuidado mais eficientes, voltadas à melhoria dos resultados em saúde e ao planejamento da rede de atenção oncológica.

2 OBJETIVOS

2.1 Geral

Explorar o potencial de diferentes bases de dados de acesso público para investigar características demográficas, tumorais e terapêuticas do câncer de mama.

2.2 Específicos

1. Descrever as características demográficas, tumorais e terapêuticas dos registros de câncer de mama.
2. Caracterizar os perfis imuno-histoquímicos e biomarcadores tumorais disponíveis.
3. Identificar e caracterizar os fatores associados à recorrência tumoral, utilizando variáveis clinicopatológicas e moleculares disponíveis.
4. Avaliar a sobrevida global e específica em cinco anos e comparar fatores prognósticos entre diferentes bases de dados.
5. Estimar e comparar as probabilidades de sobrevida global e específica em cinco anos segundo o estadiamento tumoral.
6. Caracterizar a distribuição espacial dos serviços oncológicos e os deslocamentos para tratamento no Brasil.

3 METODOLOGIA

3.1 Desenho do estudo

Foi realizado um estudo observacional, do tipo coorte retrospectiva, utilizando registros de câncer obtidos de fontes secundárias, anonimizadas e de acesso público, disponíveis na internet. Esse tipo de delineamento é amplamente empregado em pesquisas epidemiológicas e foi escolhido devido à natureza dos objetivos da pesquisa ([CHE *et al.*, 2023](#); [WANG](#); [KATTAN, 2020](#)). Embora não tenha havido envolvimento direto com seres humanos, a condução do estudo seguiu os princípios éticos. Uma consulta ao Comitê de Ética em Pesquisa com Seres Humanos da Universidade Federal de Uberlândia ([CEP-UFU](#)) foi realizada, a fim de deliberar sobre a necessidade de parecer ético, registrada sob o Certificado de Apresentação para Apreciação Ética ([CAAE](#)) 86998025.3.0000.5152 ([Anexo A](#)). Ademais, a descrição metodológica foi estruturada conforme as diretrizes do protocolo *Strengthening the Reporting of Observational Studies in Epidemiology* ([STROBE](#)) ([VON ELM *et al.*, 2007](#)), visando a melhoria da qualidade do relato ([Anexo B](#)).

3.2 Fontes de dados

A escolha dos registros baseou-se no alinhamento com os objetivos da pesquisa, adotando como critério de elegibilidade o acesso público, a disponibilidade de dados demográficos, clínico-patológicos e terapêuticos. Um instrumento de coleta de dados, do tipo *checklist*, foi elaborado para nortear a seleção das fontes (Apêndice A). Uma fonte de dados moleculares e genômicos foi incluída para ampliar a caracterização dos perfis de receptores (RE, RP, HER2). As fontes consultadas e as formas de acesso foram descritas abaixo:

- **RHC nacional:** o download foi realizado via Integrador RHC, uma interface desenvolvida pelo INCA que permitiu o acesso seletivo aos anos de registros de câncer de todos os estados brasileiros sem a necessidade de cadastro prévio. Para este estudo, foram selecionados dados disponíveis a partir do ano 2000. O endereço eletrônico de acesso foi: <<https://irhc.inca.gov.br/RHCNet/visualizaTabNetExterno.action>>
- **RHC Estadual:** a Fundação Oncocentro de São Paulo (FOSP) é a instituição responsável pela coordenação do RHC no Estado de São Paulo (SP) e disponibiliza uma interface que permitiu o download de registros de câncer por ano de diagnóstico, após um cadastro prévio no seguinte endereço eletrônico: <<https://fosp.saude.sp.gov.br/fosp/diretoria-adjunta-de-informacao-e-epidemiologia/rhc-registro-hospitalar-de-cancer/banco-de-dados-do-rhc/>>. Foram selecionados os registros disponíveis a partir de 2000. Essa fonte foi escolhida por conter informações detalhadas de seguimento.
- **Registro de Câncer de Base Populacional (RCBP):** o SEER disponibiliza diferentes produtos de pesquisa, como o SEER Research, SEER Research Plus e National Childhood Cancer Registry (NCCR), além de bases de dados especializadas (NATIONAL INSTITUTES OF HEALTH, n.d.[a]). A principal diferença entre os produtos reside na disponibilidade de variáveis específicas, que influenciam diretamente a profundidade das análises. As etapas de acesso incluíram a solicitação, a inscrição no sistema, o reconhecimento dos acordos e a limitação dos dados, bem como a instalação do software SEER*Stat. Para este estudo, foi concedida a permissão de acesso ao produto SEER Research. A lista de registros analisados foi obtida dos dados de incidência do SEER, a partir de 17 registros, submetidos em novembro de 2023, e cobriu os anos de 2000 a 2021. A versão do SEER*Stat utilizada para acessar os dados foi a 8.4.5. A solicitação de acesso foi realizada no endereço eletrônico <<https://seer.cancer.gov/data/access.html>>.
- **Registros clínicos do Molecular Taxonomy of Breast Cancer International Consortium (METABRIC):** o consórcio METABRIC é uma colaboração internacional entre o Reino Unido e o Canadá que reúne dados genômicos, transcriptômicos e clínicos sobre o câncer de mama (CURTIS *et al.*, 2012; PEREIRA *et al.*, 2016). Os dados foram acessados por meio da

plataforma cBioPortal, no seguinte endereço eletrônico <https://www.cbioportal.org/study/clinicalData?id=brca_metabric>. Não houve necessidade de cadastro prévio para download. Informações temporais detalhadas como ano de diagnóstico não estavam disponíveis na base, dessa forma, todos os registros foram incluídos, sem distinção quanto ao período de seguimento.

3.3 Participantes e critérios de inclusão

Foram considerados elegíveis os registros do sexo masculino e feminino, com idade igual ou superior a 18 anos e topografia C50³, confirmada por exame histológico do tumor primário. A inclusão de ambos os sexos justificou-se pela necessidade de representar adequadamente todos os casos registrados. A definição da idade mínima de 18 anos como critério de inclusão está alinhada à legislação brasileira, que estabelece essa faixa etária como a maioridade civil e dá autonomia plena para as decisões, inclusive aquelas relacionadas ao tratamento de saúde no país (BRASIL, 2002). É pertinente mencionar que, embora outros países possam adotar regras distintas para a idade mínima, estudos internacionais empregaram prática semelhante (LI *et al.*, 2024; QIU *et al.*, 2024). A seleção da topografia C50 baseou-se no CID-O 3ª edição, amplamente utilizado em registros de câncer (PIÑEROS *et al.*, 2017). Por fim, a confirmação microscópica por exame histológico do tumor primário visou assegurar a precisão diagnóstica e a confirmação do câncer de mama.

3.4 Critérios de exclusão

Após o download, os registros selecionados passaram por um processo de avaliação da completude e da consistência das informações. Nessa etapa, foram considerados dois pontos importantes na decisão de exclusão dos registros: (1) os objetivos da pesquisa; (2) a padronização dos dados para comparação dos fatores prognósticos entre diferentes bases de dados. Assim, foram excluídos os registros com as seguintes características:

- Ausência de informações sobre idade e sexo.
- Ausência de informações sobre o status vital (vivo/óbito). A exceção a esse critério foi o RHC nacional, que apresentou mais de 77% dos registros sem essa informação. Entretanto, considerando a importância dessa fonte de dados para os objetivos da pesquisa, optou-se por mantê-la e realizar novas análises posteriormente.

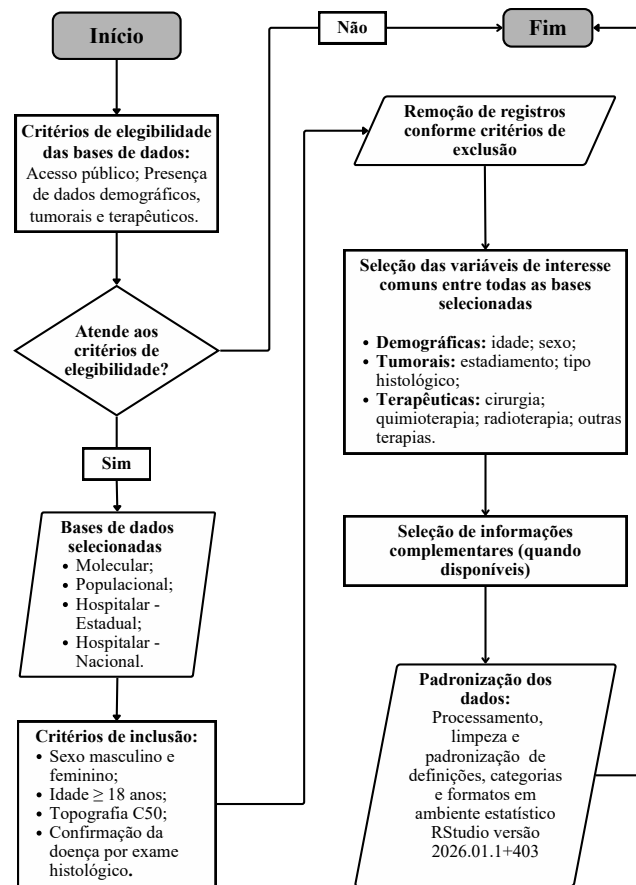
³C50 refere-se a neoplasia da mama.

- Ausência de informações sobre diagnóstico, tratamento ou tempo de seguimento.
- Registros reportados no óbito.
- Registros não analíticos⁴ (para as fontes nacionais), que, por recomendação dos registros hospitalares, devem ser excluídos das análises de sobrevida ([MINISTÉRIO DA SAÚDE, 2000](#)).
- Inconsistências nas variáveis de sobrevida, como registros com tempo de sobrevida negativo.
- Tempo de acompanhamento inferior a 30 dias. A decisão de excluir tempos de acompanhamento curtos visou reduzir o viés nas estimativas de risco nas análises de sobrevida ([NOORDZIJ *et al.*, 2013](#)).
- Registros com diagnóstico anterior a 2000 e posterior a 2015, garantindo tempo suficiente de acompanhamento para as análises de sobrevida até a data de fechamento do registro.
- O comportamento tumoral também foi padronizado, excluindo-se os casos registrados como “/0” (neoplasia benigna), “/1” (neoplasia de comportamento incerto ou desconhecido) ou “/2” (neoplasia *in situ*). Assim, os registros populacionais e hospitalares incluem exclusivamente tumores classificados como “/3” (neoplasia maligna).
- Para a base de dados moleculares, na ausência de codificação histológica [CID-O 3ª edição](#), adotou-se como critério a exclusão de casos com estadiamento clínico reportado como “0”.

3.5 Padronização

A padronização dos registros nacionais e internacionais foi planejada para permitir análises homogêneas, mesmo entre diferentes serviços de saúde e populações, preservando, contudo, a integridade da base original. O fluxo desse processo é apresentado na [Figura 1](#).

⁴São casos em que o tratamento integral (diagnóstico e conduta terapêutica inicial) não ocorreu na unidade que realiza o registro.

Figura 1 – Fluxo de seleção e padronização de dados

Fonte: Dados da pesquisa (2026).

Além do atendimento aos critérios de elegibilidade, foram considerados conjuntos de informações complementares que se articularam para uma análise mais robusta. A disponibilidade variou entre os tipos de registros, conforme apresentado na Figura 2.

Figura 2 – Comparação da disponibilidade de informações complementares nos registros analisados

	RHC Estadual	RHC Nacional	Molecular Internacional	RCBP Internacional
Grau histológico	○	○	●	●
Receptores (RE, RP, HER2)	○	○	●	●
Subtipos PAM50	○	○	●	○
Status menopausal	○	○	●	○
Status de recorrência do tumor	●	○	●	○
Localização domicílio/tratamento	●	●	○	○
Data do diagnóstico	●	●	○	○
Data do óbito	●	●	○	○
Data do último contato (follow-up)	●	○	○	○
Tempo de sobrevida	○	○	●	●

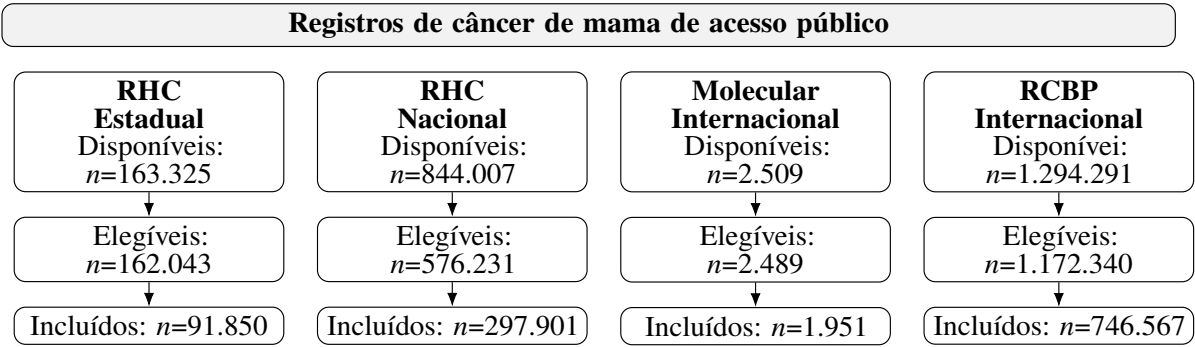
Nota: Círculo preenchido (●) indica que a informação estava disponível; Círculo vazio (○) indica ausência de informação.

Fonte: Dados da pesquisa (2026).

Como parte do processo de padronização, foram criadas variáveis derivadas, incluindo informações sobre a distância, obtidas através do identificador numérico oficial, atribuído pelo Instituto Brasileiro de Geografia e Estatística (IBGE), ao município de residência e ao município de tratamento (BRASIL, 2025). Essa etapa permitiu categorizar os registros hospitalares brasileiros em faixas de distância percorrida e investigar desigualdades geográficas no acesso a tratamentos.

A consistência da codificação para a classificação histológica foi verificada adotando-se como referência os códigos de morfologia especificados na 8ª edição da AJCC para todas as fontes (HORTOBAGYI *et al.*, 2017). Adicionalmente, o estadiamento codificado como “X”, “9”, “99”, ou qualquer outro código não padronizado pela AJCC⁵ foi mantido como ausente para análise de sensibilidade posterior. Para o RHC Estadual, adotou-se a edição do TNM informada na própria base de dados (diagnóstico até 2005: TNM 5ª edição; de 2006 a 2013: TNM 6ª edição; a partir de 2014: TNM 7ª edição). No RCBP Internacional, adotou-se a 6ª edição da AJCC (diagnósticos de 1988 a 2015). Os detalhes do número de registros incluídos após o processo de padronização são apresentados na Figura 3.

Figura 3 – Detalhamento dos registros incluídos na pesquisa nas fontes analisadas.



Nota: Disponíveis: número de registros na base de dados. Elegíveis: registros que atenderam aos critérios de inclusão (sexo masculino ou feminino; idade maior ou igual a 18 anos; topografia C50; confirmação histológica do tumor primário). Incluídos: registros elegíveis selecionados para compor a amostra. As exclusões específicas decorrentes de dados faltantes são detalhadas nos fluxogramas de cada análise.

Fonte: Dados da pesquisa (2026).

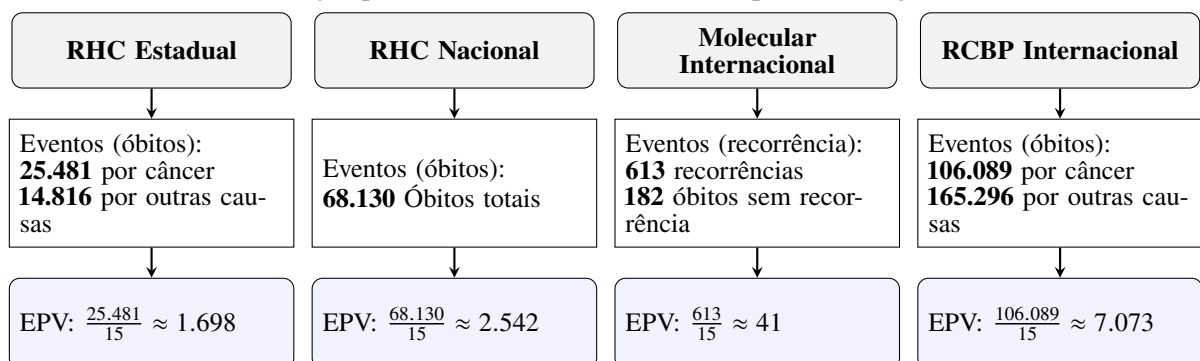
3.6 Tamanho da amostra

Para este estudo, o tamanho da amostra foi definido pela totalidade de registros qualificados (incluídos) nas quatro fontes de dados, totalizando 1.138.269 registros de câncer de mama. A

⁵O estadiamento anatômico da AJCC utiliza algarismos romanos (I, II, III e IV) e letras (A, B e C) para indicar o avanço da doença, sendo os tumores *in situ* indicados com numeral arábico zero (0). Os registros de câncer utilizaram códigos como “X”, “9” ou “99” para indicar que o estadiamento não foi informado ou não pôde ser avaliado.

adequação da amostra para as análises de risco multivariadas foi validada pelo critério de Eventos por Variável (EPV), visando assegurar a estabilidade das estimativas. Considerando o plano analítico com até 15 variáveis independentes, adotou-se o limiar mínimo de 10 EPV para a modelagem multivariada (AUSTIN; ALLIGNOL; FINE, 2017), garantindo que todas as fontes selecionadas apresentassem eventos suficientes. O detalhamento do número de eventos para cada fonte de dados analisada é apresentado na Figura 4.

Figura 4 – Avaliação preliminar do número de eventos para modelagem multivariada.



Fonte: Dados da pesquisa (2026).

3.7 Análise Estatística

A estatística descritiva foi utilizada para caracterizar os diferentes tipos de registros avaliados no estudo. Variáveis contínuas foram expressas em medianas e intervalos interquartílicos (IIQ), enquanto variáveis categóricas foram descritas por frequências absolutas (n) e relativas (%).

Para a análise da recorrência tumoral, foram utilizados modelos de riscos competitivos de Fine–Gray para estimar as razões de risco de subdistribuição (sHR), considerando o óbito sem recorrência como evento competidor⁶. As probabilidades cumulativas de recorrência foram estimadas por meio de Função de Incidência Cumulativa (CIF) e comparadas entre os subtipos moleculares através do Teste de Gray. A seleção dos preditores inseridos nos modelos multivariados foi estruturada *a priori* com base em pressupostos teóricos mapeados por *Directed Acyclic Graph* (DAG), permitindo a identificação de potenciais fontes de viés, como confundimento e ajustes inadequados. A ausência de multicolinearidade linear foi confirmada pelo Fator de Inflação da Variância Generalizado (*Generalized Variance Inflation Factor* – GVIF) ajustado

⁶Evento competidor é um acontecimento alternativo que impede que o desfecho primário de interesse venha a acontecer e por isso precisa ser tratado de forma diferente da censura comum na análise de sobrevivência.

(FOX; MONETTE, 1992). Como o modelo de Fine-Gray não dispõe de procedimento equivalente ao teste de Schoenfeld para avaliação da proporcionalidade dos riscos, esse pressuposto foi avaliado indiretamente por meio de um modelo de Cox causa-específico, utilizando o teste de Grambsch–Therneau (GRAMBSCH; THERNEAU, 1994). Adicionalmente, foram inspecionados os resíduos dos modelos de Fine-Gray para avaliação gráfica. Quando identificada a violação da proporcionalidade⁷, foram reportadas estimativas médias de efeito (*Average Hazard Ratios*), fundamentadas nos princípios descritos na literatura (BELLERA *et al.*, 2010; SCHEMPER; WAKOUNIG; HEINZE, 2009; FINE; GRAY, 1999). Os procedimentos de diagnóstico e validação do modelo de recorrência são apresentados no Apêndice B.

Para a análise de sobrevida, a sobrevida global e a sobrevida específica por câncer de mama foram estimadas por meio de modelos de riscos proporcionais de Cox e modelos de Cox causa-específicos. Curvas de sobrevivência ajustadas, preditas a partir dos modelos multivariados de Cox foram geradas para ilustrar a sobrevida global e a sobrevida específica de 5 anos. A seleção dos preditores inseridos nos modelos multivariados também foi mapeada por DAG e a ausência de multicolinearidade⁸ foi confirmada pelo GVIF ajustado. A proporcionalidade dos riscos foi avaliada pelo teste de Grambsch–Therneau e resíduos de Schoenfeld (GRAMBSCH; THERNEAU, 1994). Quando o pressuposto foi violado, foram reportadas estimativas médias de efeito. Os procedimentos de diagnóstico e validação do modelo de sobrevida global e específica são apresentados no Apêndice C.

Intervalos de Confiança de 95% (IC 95%) foram adotados em todas as estimativas e valores de $\alpha < 0,05$ foram considerados estatisticamente significativos. Todas as análises foram conduzidas no ambiente estatístico RStudio, versão 2026.01.1+403, utilizando os pacotes estatísticos *tidyverse* (para manipulação dos dados e geração dos gráficos), *survival* (para os modelos de Cox), *cmprsk* (para análises de Fine-Gray e teste de Gray), *geobr* (para a extração dos dados geográficos) e *osrm* (para o cálculo de rotas reais de deslocamento).

⁷O pressuposto de riscos proporcionais assume que o efeito de uma variável explicativa sobre o risco não varia durante o período de seguimento. Quando esse pressuposto é violado, a interpretação dos coeficientes do modelo torna-se limitada, exigindo estratégias alternativas, como o uso de modelos com termos dependentes do tempo, estratificação por variáveis ou a apresentação de estimativas médias de efeito.

⁸A multicolinearidade ocorre quando há forte correlação entre as variáveis explicativas, o que aumenta a variância das estimativas dos coeficientes e pode dificultar a identificação do efeito individual de cada variável.

4 RESULTADOS

4.1 Caracterização dos perfis imuno-histoquímicos e genômicos

Para alcançar o objetivo de caracterizar os perfis imuno-histoquímicos e genômicos dos tumores, optou-se pela utilização exclusiva do Registro Molecular Internacional. Essa escolha justifica-se pela qualidade das informações moleculares fornecidas por essa fonte, incluindo a classificação genômica por subtipos intrínsecos, variável não disponível nas demais fontes de dados analisadas (Figura 2). A extração seletiva de marcadores moleculares por **IHQ** e pelos subtipos intrínsecos classificados pelo **PAM50** resultou em 1.951 registros analisados. Os resultados são apresentados nas Tabelas 1 e 2.

Tabela 1 – Caracterização molecular pela imuno-histoquímica

Marcadores por IHQ	<i>n</i>	%
Receptor de Estrogênio		
Positivo	1.488	76,27
Negativo	463	23,73
Receptor de Progesterona		
Positivo	1.025	52,54
Negativo	924	47,36
Ausente	2	0,10
HER2		
Positivo	243	12,46
Negativo	1.707	87,49
Ausente	1	0,05

Nota: IHQ: imuno-histoquímica; N=1.951

Fonte: Dados da pesquisa (2026).

Tabela 2 – Caracterização molecular por subtipos intrínsecos classificados pelo PAM50

Subtipos intrínsecos	<i>n</i>	%
Luminal A	694	35,57
Luminal B	468	23,99
HER2 enriquecido	221	11,33
<i>Claudin-low</i>	211	10,81
<i>Basal-like</i>	205	10,51
<i>Normal-like</i>	145	7,43
Indeterminado	6	0,31
Ausente	1	0,05

Nota: PAM50: *Prediction Analysis of Microarray 50*; *N*=1.951.

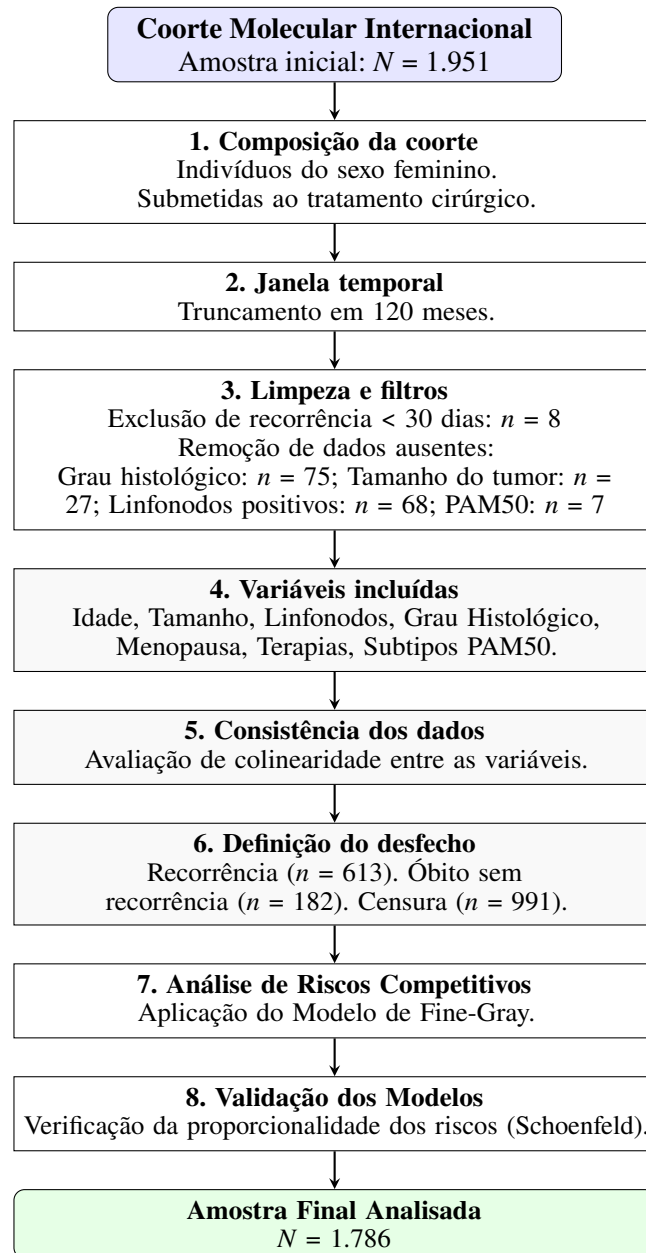
Fonte: Dados da pesquisa (2026).

A análise do registro Molecular Internacional revelou a distribuição das variáveis moleculares fundamentais por meio da classificação fenotípica (IHQ) e genômica (PAM50). Conforme apresentado nas Tabelas 1 e 2, a maioria dos registros foi classificada com sucesso. Observou-se uma positividade de 76,27% para o receptor de estrogênio e 52,54% para o receptor de progesterona, enquanto o HER2 apresentou-se positivo em 12,46% dos registros. Quanto à classificação genômica, os subtipos Luminal A (35,57%) e Luminal B (23,99%) foram os mais frequentes. Em conjunto, esses resultados evidenciam a predominância de tumores luminais na coorte analisada.

4.2 Caracterização clínica e terapêutica da coorte molecular segundo os subtipos intrínsecos

Para as análises subsequentes, os dados foram submetidos a procedimentos de pré-processamento para garantir a consistência das informações e a adequação às análises de recorrência. A exclusão dos registros com recorrência inferior a 30 dias foi adotada, visto que tais eventos não refletem a biologia da recidiva tumoral após controle inicial, mas sim a persistência da doença não detectada ou não ressecada adequadamente no ato operatório. O truncamento em 120 meses buscou refletir a prática clínica e a relevância desse intervalo para a monitorização da recorrência. Os detalhes do pré-processamento são apresentados na Figura 5.

Figura 5 – Fluxo de pré-processamento dos dados para avaliação da recorrência.



Nota: N indica o tamanho da amostra total; n indica o número de observações. A soma das exclusões individuais na Etapa 3 ($n = 185$) difere do total de indivíduos excluídos ($n = 165$) devido à sobreposição de critérios, visto que alguns registros apresentavam dados ausentes em mais de uma variável simultaneamente.

Fonte: Dados da pesquisa (2026).

A coorte foi composta por 1.786 registros, nos quais se observou variação etária entre os diferentes subtipos intrínsecos. Mulheres com subtipos Luminal A e B apresentaram medianas de idade mais elevadas (63 e 66 anos, respectivamente), enquanto o subtipo *Basal-like* concentrou pacientes mais jovens, com mediana de 55 anos (IIQ: 44-66). Este perfil etário refletiu o status menopausal, em que o subtipo Luminal B apresentou a maior proporção de pacientes na pós-menopausa (88,99%), em contraste com o subtipo *Basal-like*, que registrou o maior percentual

de mulheres na pré-menopausa (38,34%).

No que se refere ao tamanho do tumor, observou-se que os subtipos **HER2** enriquecido, Luminal B e *Basal-like* apresentaram maior dimensão ao diagnóstico, com maior proporção de registros classificados na categoria de 21 a 50 milímetros (58,25%; 58,12% e 56,48%, respectivamente). Em contraste, o subtipo *Claudin-low* destacou-se pela maior frequência de tumores pequenos, com 54,59% dos registros medindo até 20 milímetros.

A ausência de comprometimento linfonodal foi mais frequente nos subtipos *Normal-like* (58,40%) e Luminal A (57,50%). Em contrapartida, os subtipos **HER2** enriquecido, *Claudin-low* e *Basal-like* apresentaram maior prevalência de comprometimento linfonodal extenso, com frequências elevadas na categoria de 10 ou mais linfonodos (8,25%; 8,11% e 7,77%, respectivamente).

Maior frequência de alto grau histológico foi observada nos subtipos *Basal-like* e **HER2** enriquecido (91,71% e 74,76%, respectivamente). Em oposição, o subtipo Luminal A exibiu o perfil mais indolente, com 74,22% dos registros classificados como grau histológico baixo/intermediário.

O manejo terapêutico diferiu entre os subtipos moleculares intrínsecos. A quimioterapia foi predominantemente administrada nos subtipos *Basal-like* (54,40%), enquanto nos grupos Luminais A e B essa modalidade terapêutica não ultrapassou 10%. Por outro lado, a hormonioterapia foi a base do tratamento para os subtipos Luminal B (80,09%) e Luminal A (68,28%). O uso de radioterapia foi consistente entre os subtipos, com frequências variando entre 53,60% no *Normal-like* e 70,81% no *Claudin-low*.

Ao final de 10 anos de seguimento, o subtipo **HER2** enriquecido apresentou a maior proporção de recorrência (49,03%), seguido pelo Luminal B (42,56%) e *Basal-like* (41,45%). Observou-se uma elevada distribuição de óbitos sem recorrência atuando como evento competidor nos subtipos **HER2** enriquecido (18,93%) e Luminal A (17,65%). Mais detalhes sobre as características da coorte são apresentados na Tabela 3.

Tabela 3 – Caracterização clínica e terapêutica da coorte molecular segundo os subtipos intrínsecos classificados pelo PAM50

Variável	<i>Basal-like</i> n (%)	<i>Claudin-low</i> n (%)	HER2 enriquecido n (%)	Luminal A n (%)	Luminal B n (%)	<i>Normal-like</i> n (%)
Idade (anos)						
Mediana (IIQ)	55 (44–66)	58 (48–67)	59 (50–69)	63 (53–72)	66 (58–74)	57 (46–66)
Tamanho do tumor						
Até 20 milímetros	69 (35,75%)	101 (54,59%)	70 (33,98%)	319 (49,84%)	161 (36,84%)	57 (45,60%)
21 a 50 milímetros	109 (56,48%)	76 (41,08%)	120 (58,25%)	303 (47,34%)	254 (58,12%)	59 (47,20%)
>50 milímetros	15 (7,77%)	8 (4,32%)	16 (7,77%)	18 (2,81%)	22 (5,03%)	9 (7,20%)
Linfonodos positivos						
Ausentes	90 (46,63%)	96 (51,89%)	85 (41,26%)	368 (57,50%)	213 (48,74%)	73 (58,40%)
1 a 3 linfonodos	75 (38,86%)	53 (28,65%)	73 (35,44%)	190 (29,69%)	148 (33,87%)	34 (27,20%)
4 a 9 linfonodos	13 (6,74%)	21 (11,35%)	31 (15,05%)	63 (9,84%)	47 (10,76%)	15 (12,00%)
10 ou mais linfonodos	15 (7,77%)	15 (8,11%)	17 (8,25%)	19 (2,97%)	29 (6,64%)	3 (2,40%)
Grau histológico						
Baixo ou Intermediário	16 (8,29%)	59 (31,89%)	52 (25,24%)	475 (74,22%)	193 (44,17%)	82 (65,60%)
Alto	177 (91,71%)	126 (68,11%)	154 (74,76%)	165 (25,78%)	244 (55,83%)	43 (34,40%)
Quimioterapia						
Sim	105 (54,40%)	64 (34,59%)	86 (41,75%)	54 (8,44%)	43 (9,84%)	29 (23,20%)
Não	88 (45,60%)	121 (65,41%)	120 (58,25%)	586 (91,56%)	394 (90,16%)	96 (76,80%)
Hormonioterapia						
Sim	57 (29,53%)	94 (50,81%)	90 (43,69%)	437 (68,28%)	350 (80,09%)	79 (63,20%)
Não	136 (70,47%)	91 (49,19%)	116 (56,31%)	203 (31,72%)	87 (19,91%)	46 (36,80%)
Radioterapia						
Sim	133 (68,91%)	131 (70,81%)	127 (61,65%)	354 (55,31%)	279 (63,84%)	67 (53,60%)
Não	60 (31,09%)	54 (29,19%)	79 (38,35%)	286 (44,69%)	158 (36,16%)	58 (46,40%)
Status menopausal						
Pós-menopausa	119 (61,66%)	129 (69,73%)	156 (75,73%)	518 (80,94%)	389 (88,99%)	81 (64,80%)
Pré-menopausa	74 (38,34%)	56 (30,27%)	50 (24,27%)	122 (19,06%)	48 (11,01%)	44 (35,20%)
Desfecho aos 120 meses						
Censura (sem recorrência)	90 (46,63%)	123 (66,49%)	66 (32,04%)	377 (58,91%)	185 (42,33%)	63 (50,40%)
Recorrência	80 (41,45%)	50 (27,03%)	101 (49,03%)	150 (23,44%)	186 (42,56%)	46 (36,80%)
Óbito sem recorrência	23 (11,92%)	12 (6,48%)	39 (18,93%)	113 (17,65%)	66 (15,11%)	16 (12,80%)

Notas: IIQ: Intervalo Interquartilico (Q1–Q3); *Basal-like*: N=193; *Claudin-low*: N=185; HER2 enriquecido: N=206; Luminal A: N=640; Luminal B: N=437; Normal-like: N=125.

Fonte: Dados da pesquisa (2026).

4.3 Fatores prognósticos associados ao risco de recorrência do câncer de mama

Considerando a relevância dos subtipos moleculares intrínsecos no prognóstico da doença, um modelo de riscos competitivos foi elaborado para incluir e avaliar o efeito dos subtipos intrínsecos sobre a recorrência do câncer de mama em mulheres ao longo de 120 meses. A seleção dos preditores inseridos no modelo foi fundamentada no conhecimento biológico e clínico descrito na literatura (ALLISON, 2021; RUEDA *et al.*, 2019) e estruturada por meio de um grafo acíclico direcionado (DAG). Ressalta-se que, embora o delineamento observacional do estudo não permita a confirmação inequívoca de nexos causais, a estrutura causal do grafo serviu como ferramenta metodológica para o mapeamento dos pressupostos teóricos e a identificação de potenciais fontes de viés, como confundimento e ajustes inadequados (TEXTOR *et al.*, 2017). Essa representação gráfica do modelo é apresentada no Apêndice B (Figura B1).

A multicolinearidade das variáveis inseridas no modelo (idade, tamanho do tumor, linfonodos, grau histológico e subtipos moleculares) foi avaliada por meio do fator de inflação de variância generalizado – *Generalized Variance Inflation Factor* (GVIF). O GVIF ajustado máximo observado foi de 1,071 (Tabela B1), indicando ausência de multicolinearidade relevante entre as variáveis (O'BRIEN, 2007). Tratamentos adjuvantes foram considerados mediadores e, portanto, não incluídos no conjunto de ajuste.

O diagnóstico de ajuste do modelo de Fine–Gray combinou a avaliação gráfica dos resíduos (Figura B2) e o teste de correlação de Grambsch e Therneau, derivado de um modelo Cox causa-específico (Tabela B2), como método de avaliação indireta da proporcionalidade dos riscos. O teste indicou violação nas variáveis idade ($p = 0,017$), grau histológico ($p < 0,001$) e subtipo molecular ($p < 0,001$), sugerindo que o efeito dessas variáveis sobre a recorrência do câncer de mama não é constante ao longo do tempo (Tabela B3). Diante disso, o modelo foi segmentado em janelas temporais. Os detalhes desse diagnóstico e da avaliação dos pressupostos do modelo estão reunidos no Apêndice B.

Diante dessa natureza dinâmica, as razões de subdistribuição de riscos (sHR) apresentadas nos seguimentos completo (0–120 meses) e precoce (0–60 meses) foram interpretadas como estimativas médias de efeito (*Average Hazard Ratios*), fundamentadas nos princípios descritos na literatura (SCHEMPER; WAKOUNIG; HEINZE, 2009; FINE; GRAY, 1999). Essa abordagem permitiu quantificar o impacto predominante de cada variável, medido como a média ponderada do efeito ao longo do tempo. Assim, para facilitar a distinção entre os efeitos médios e constantes, adotou-se a notação sHR_{médio} para indicar as janelas temporais em que a premissa da proporcionalidade foi violada.

Os resultados das análises indicaram um efeito protetor incremental da idade no modelo univariado do seguimento precoce (0–60 meses). Cada ano a mais de vida reduziu o risco médio de recorrência em 1,3% (sHR_{médio}: 0,987; IC 95%: 0,980–0,995; $p < 0,001$). No entanto, esse

efeito foi atenuado quando ajustado pelos subtipos PAM50, grau histológico, tamanho do tumor e linfonodos positivos. No modelo multivariado (Tabela 5), a redução no risco médio associado à idade passou de 1,3% para 0,8% no seguimento precoce (sHR_{médio}: 0,992) e perdeu o efeito no seguimento tardio (sHR: 0,998; IC 95%: 0,988–1,008; $p = 0,690$).

O tamanho do tumor apresentou associação independente com a recorrência. No seguimento completo da análise multivariada (Tabela 5), cada milímetro adicional do tumor elevou o risco médio de recorrência em 0,9% (sHR_{médio}: 1,009; IC 95%: 1,005–1,014; $p < 0,001$). O risco manteve-se praticamente inalterado no seguimento precoce (sHR_{médio}: 1,010; IC 95%: 1,005–1,015; $p < 0,001$) e no seguimento tardio (sHR: 1,009; IC 95%: 1,001–1,018; $p = 0,030$).

O número de linfonodos positivos elevou o risco médio ajustado de recorrência em 5,3% (sHR_{médio}: 1,053; IC 95%: 1,032–1,076; $p < 0,001$) no seguimento precoce (Tabela 5). Diferente de outros preditores, os linfonodos demonstraram grande consistência temporal, mantendo magnitude semelhante no seguimento tardio (61–120 meses), com incremento de 5,6% (sHR: 1,056; IC 95%: 1,029–1,083; $p < 0,001$) por linfonodo positivo no risco de recorrência, quando comparados a indivíduos sem comprometimento linfonodal.

O grau histológico apresentou associação com o risco médio de recorrência precoce, no entanto, perdeu impacto no prognóstico ao longo do tempo. Na análise univariada (Tabela 4), pacientes com tumores Grau 3 apresentaram um aumento de 216,9% no risco médio de recorrência em relação ao Grau 1 nos primeiros 60 meses (sHR_{médio}: 3,169; IC 95%: 2,002–5,017; $p < 0,001$). No entanto, no seguimento tardio (61–120 meses), essa associação perdeu significância estatística ($p = 0,180$). Ao observar o efeito direto no modelo multivariado (Tabela 5), o Grau 3 manteve-se como um preditor independente de risco na janela precoce, embora com magnitude reduzida após ajuste, em que o risco foi de 77,6% (sHR_{médio}: 1,776; IC 95%: 1,115–2,828; $p = 0,016$). Um comportamento interessante foi observado nos tumores de Grau 2, mantendo-se como fator de risco significativo no seguimento completo do modelo multivariado (sHR_{médio}: 1,429; IC 95%: 1,039–1,966; $p = 0,028$) e tangenciando a significância no seguimento tardio (sHR: 1,550; IC 95%: 0,983–2,446; $p = 0,059$).

A análise dos subtipos intrínsecos revelou padrões distintos de risco de recorrência ao longo do período observado. Na análise multivariada (Tabela 5), no seguimento precoce (0–60 meses), todos os subtipos, com exceção do *Claudin-low* ($p = 0,250$), apresentaram associação estatisticamente significativa com o risco de recorrência em 60 meses (5 anos). Em contrapartida, no seguimento tardio (61–120 meses), apenas o subtipo *Basal-like* manteve significância estatística ($p < 0,001$), confirmando que o efeito dos subtipos sobre a recorrência não é constante ao longo do tempo.

Para o subtipo Luminal B, observou-se, após ajuste (Tabela 5), que o risco de recorrência do tumor concentra-se nos primeiros 5 anos de seguimento, com um aumento de 82,7% no risco médio em comparação ao subtipo Luminal A (sHR_{médio}: 1,827; IC 95%: 1,386–2,407; $p < 0,001$).

O subtipo **HER2** enriquecido também demonstrou risco elevado após ajuste no período inicial (0–60 meses), com aumento no risco médio de 126,7% (sHR_{médio}: 2,267; IC 95% 1,642–3,130; $p < 0,001$), consolidando-se, juntamente com o subtipo *Basal-like*, como os perfis de maior risco de recorrência nos primeiros 5 anos.

O comportamento do subtipo *Basal-like* foi heterogêneo ao longo dos períodos de seguimento. Na análise multivariada (Tabela 5), observou-se que, no período precoce (0–60 meses), esse subtipo apresentou um risco médio elevado, de 136,6% em relação ao Luminal A (sHR_{médio}: 2,366; IC 95%: 1,679–3,334; $p < 0,001$). No seguimento tardio, observou-se uma inversão das estimativas, sugerindo uma redução de 73,6% no risco de recidiva após 5 anos (sHR: 0,264; IC 95%: 0,121–0,577; $p < 0,001$).

Em relação ao subtipo *Claudin-low*, a análise multivariada indicou a ausência de diferença significativa no risco de recorrência em comparação ao subtipo Luminal A, tanto no seguimento precoce quanto no tardio ($p = 0,250$ e $p = 0,200$, respectivamente). Esse resultado sugere que, nesta coorte, após ajuste, o comportamento de recorrência do subtipo *Claudin-low* assemelha-se ao do subtipo Luminal A.

Por fim, o subtipo *Normal-Like* manteve-se estatisticamente significativo no seguimento precoce após ajuste, contribuindo com um aumento de 51,3% no risco médio de recorrência precoce (sHR_{médio}: 1,513; IC 95%: 1,006–2,277; $p = 0,047$), perdendo significância após o quinto ano de seguimento ($p = 0,190$). As Tabelas 4 e 5 consolidam as análises univariadas e multivariadas, respectivamente, visando estimar os efeitos total e direto dos preditores sobre a recorrência.

Tabela 4 – Análise univariada Fine-Gray para o efeito total dos preditores sobre a recorrência

Variável	Seguimento Completo (0–120 meses)		Seguimento Precoce (0–60 meses)		Seguimento Tardio (61–120 meses)	
	sHR ^(a) (IC 95%)	<i>p</i>	sHR ^(a) (IC 95%)	<i>p</i>	sHR (IC 95%)	<i>p</i>
Variáveis contínuas^(b)						
Idade (anos)	0,991 (0,985–0,997)	0,003	0,987 (0,980–0,995)	<0,001	1,003 (0,994–1,012)	0,510
Tamanho do tumor (mm)	1,014 (1,009–1,019)	<0,001	1,017 (1,012–1,022)	<0,001	1,008 (1,001–1,016)	0,026
Linfonodos positivos	1,064 (1,043–1,087)	<0,001	1,070 (1,048–1,093)	<0,001	1,061 (1,035–1,088)	<0,001
Grau Histológico						
Grau 1 (Referência)						
Grau 2	1,650 (1,194–2,280)	0,002	1,662 (1,033–2,673)	0,036	1,670 (1,059–2,633)	0,027
Grau 3	2,250 (1,636–3,093)	<0,001	3,169 (2,002–5,017)	<0,001	1,370 (0,865–2,171)	0,180
Subtipo PAM50						
Luminal A (Referência)						
Luminal B	1,726 (1,435–2,076)	<0,001	2,236 (1,719–2,909)	<0,001	1,323 (1,001–1,748)	0,049
HER2 enriquecido	2,099 (1,655–2,662)	<0,001	3,270 (2,427–4,405)	<0,001	1,051 (0,686–1,610)	0,820
Basal-like	1,714 (1,303–2,253)	<0,001	3,404 (2,493–4,647)	<0,001	0,295 (0,145–0,599)	<0,001
Claudin-low	1,120 (0,845–1,484)	0,430	1,533 (1,048–2,245)	0,028	0,745 (0,485–1,144)	0,180
Normal-like	1,512 (1,144–1,998)	0,004	1,714 (1,146–2,561)	0,009	1,406 (0,930–2,126)	0,110

Nota: sHR: Razão de riscos de subdistribuição (*sub-distribution Hazard Ratio*); IC 95%: intervalo de confiança de 95%; *p*: significância estatística. Seguimento Completo e Precoce: *n* = 1.786; recorrências = 447. Seguimento Tardio: *n* = 1.212; recorrências = 166. Os registros analisados no seguimento tardio (61–120 meses) estavam livres de recorrência nos primeiros 60 meses.

^(a) Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade verificada no período.

^(b) Para variáveis contínuas, o sHR representa o risco associado ao incremento de uma unidade (1 ano, 1 mm ou 1 linfonodo).

Fonte: Dados da pesquisa (2026).

Tabela 5 – Análise multivariada Fine-Gray para o efeito ajustado dos preditores sobre a recorrência

Variável	Seguimento Completo (0–120 meses)		Seguimento Precoce (0–60 meses)		Seguimento Tardio (61–120 meses)	
	sHR ^(a)		sHR ^(a)		sHR	
	(IC 95%)	<i>p</i>	(IC 95%)	<i>p</i>	(IC 95%)	<i>p</i>
Variáveis contínuas^(b)						
Idade (anos)	0,992 (0,986–0,998)	0,014	0,992 (0,984–1,000)	0,056	0,998 (0,988–1,008)	0,690
Tamanho do tumor (mm)	1,009 (1,005–1,014)	<0,001	1,010 (1,005–1,015)	<0,001	1,009 (1,001–1,018)	0,030
Linfonodos positivos	1,052 (1,031–1,073)	<0,001	1,053 (1,032–1,076)	<0,001	1,056 (1,029–1,083)	<0,001
Grau Histológico						
Grau 1 (Referência)						
Grau 2	1,429 (1,039–1,966)	0,028	1,357 (0,853–2,160)	0,200	1,550 (0,983–2,446)	0,059
Grau 3	1,576 (1,136–2,186)	0,006	1,776 (1,115–2,828)	0,016	1,371 (0,840–2,237)	0,210
Subtipo PAM50						
Luminal A (Referência)						
Luminal B	1,502 (1,237–1,825)	<0,001	1,827 (1,386–2,407)	<0,001	1,220 (0,904–1,646)	0,190
HER2 enriquecido	1,607 (1,238–2,085)	<0,001	2,267 (1,642–3,130)	<0,001	1,004 (0,636–1,583)	0,990
Basal-like	1,307 (0,966–1,769)	0,082	2,366 (1,679–3,334)	<0,001	0,264 (0,121–0,577) ^(c)	<0,001
Claudin-low	0,980 (0,733–1,310)	0,890	1,250 (0,854–1,829)	0,250	0,744 (0,471–1,174)	0,200
Normal-like	1,347 (1,018–1,780)	0,037	1,513 (1,006–2,277)	0,047	1,322 (0,873–2,002)	0,190

Nota: sHR: Razão de riscos de subdistribuição (*sub-distribution Hazard Ratio*); IC 95%: intervalo de confiança de 95%; *p*: significância estatística. Seguimento Completo e Precoce: *n* = 1.786; recorrências = 447. Seguimento Tardio: *n* = 1.212; recorrências = 166. Os registros analisados no seguimento tardio (61–120 meses) estavam livres de recorrência nos primeiros 60 meses.

^(a) Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade verificada no período.

^(b) Para variáveis contínuas, o sHR representa o risco associado ao incremento de uma unidade (1 ano, 1 mm ou 1 linfonodo).

^(c) A instabilidade das estimativas tardias no subtipo *Basal-like* refletiu a presença simultânea de poucos eventos tardios (*n* = 4) e de óbitos sem recorrência (*n* = 18), que atuaram como eventos competidores.

Fonte: Dados da pesquisa (2026).

4.4 Análise da incidência cumulativa

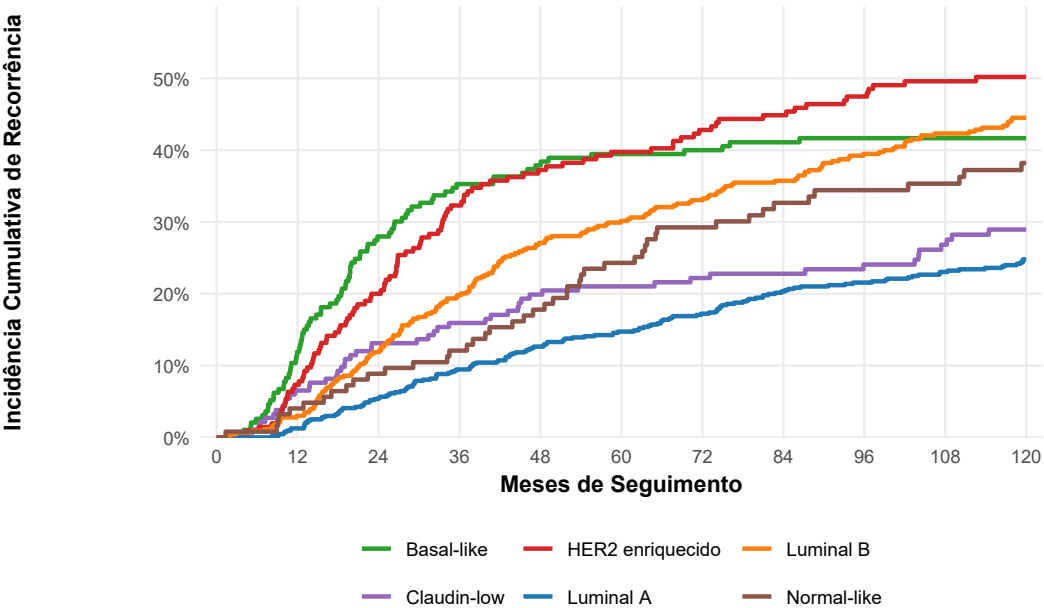
Uma vez identificados os preditores independentes do risco de recorrência, foi estimada a incidência cumulativa por subtipos intrínsecos, utilizando funções de incidência cumulativa⁹.

⁹Incidência cumulativa estima a probabilidade de um evento específico ocorrer, no caso, a recorrência do tumor. Diferente de outras medidas, a incidência cumulativa "acumula" os eventos à medida que o tempo passa.

Essa análise permitiu uma compreensão mais refinada da força do subtipo intrínseco no desfecho estudado, complementando as estimativas de risco obtidas nas análises multivariadas.

A Figura 6 apresenta as curvas de incidência cumulativa total dos subtipos intrínsecos ao longo de 120 meses de seguimento. Na análise visual das curvas, observa-se a dinâmica temporal dos riscos, com cruzamentos e mudanças de inclinação das curvas, justificando a estratégia de análise por seguimento precoce e tardio.

Figura 6 – Incidência cumulativa de recorrência por subtipo intrínseco classificado pelo PAM50



At Risk											
Basal-like	193	169	136	120	115	110	105	96	92	86	76
Claudin-low	185	171	156	148	139	135	128	119	109	97	87
HER2 enriquecido	206	188	159	133	121	113	104	97	88	76	67
Luminal A	640	630	595	557	526	493	453	416	383	350	310
Luminal B	437	416	372	334	293	275	251	228	202	178	157
Normal-like	125	118	110	103	95	86	79	69	64	56	48

Nota: Teste de Gray: $p < 0,001$. O Teste de Gray é o equivalente ao teste de Log-Rank, mas desenhado especificamente para análises que envolvem riscos competitivos. As curvas mostram a incidência cumulativa de recorrência ao longo de 120 meses, considerando óbito por outras causas como evento competidor. A tabela abaixo do gráfico indica o número de indivíduos em risco em cada intervalo de tempo.

Fonte: Dados da pesquisa (2026).

Na Tabela 6 são apresentadas as estimativas pontuais e os intervalos de confiança de 95% aos 60 e 120 meses para facilitar a interpretação das curvas. Os resultados confirmam numericamente as trajetórias observadas na Figura 6.

Tabela 6 – Incidência cumulativa de recorrência do câncer de mama por subtipo intrínseco

Subtipo intrínseco	Aos 60 meses		Aos 120 meses	
	Incidência cumulativa (IC 95%)	n em risco	Incidência cumulativa (IC 95%)	n em risco
Luminal A	14,74% (12,10–17,62)	493	24,78% (21,39–28,31)	310
Luminal B	30,16% (25,87–34,56)	275	44,52% (39,66–49,26)	157
HER2 enriquecido	39,76% (33,00–46,42)	113	50,20% (43,04–56,94)	67
<i>Basal-like</i>	39,46% (32,53–46,31)	110	41,68% (34,62–48,57)	76
<i>Claudin-low</i>	21,03% (15,41–27,26)	135	28,95% (22,28–35,93)	87
<i>Normal-like</i>	24,31% (17,12–32,19)	86	38,21% (29,46–46,90)	48

Nota: IC 95%: intervalo de confiança de 95%.

Fonte: Dados da pesquisa (2026).

Para o subtipo Luminal A, as curvas exibiram uma incidência cumulativa de recorrência de 14,74% aos 60 meses (IC 95%: 12,10–17,62), atingindo 24,78% aos 120 meses (IC 95%: 21,39–28,31). Em contraste, o subtipo [HER2](#) enriquecido apresentou as maiores incidências cumulativas no período analisado, com 39,76% aos 60 meses (IC 95%: 33,00–46,42) e atingindo 50,20% (IC 95%: 43,04–56,94) aos 120 meses.

O subtipo Luminal B apresentou uma incidência cumulativa de recorrência de 44,52% (IC 95%: 39,66–49,26) ao final de 120 meses de seguimento. Das 186 recorrências observadas nesse subtipo, 129 (69,4%) ocorreram no seguimento precoce (Tabela [B4](#)), sendo consistente com os resultados obtidos na análise multivariada (Tabela [5](#)).

O platô observado na curva correspondente ao subtipo *Basal-like* após 60 meses foi compatível com a baixa ocorrência de recorrências na janela tardia. Conforme a Tabela [B4](#), apenas quatro recorrências foram registradas entre 61 e 120 meses entre 110 indivíduos que permaneceram em risco nesse período. Consequentemente, a incidência cumulativa aumentou discretamente de 39,46% aos 60 meses para 41,68% (IC 95%: 34,62–48,57) aos 120 meses.

Quanto aos subtipos *Claudin-low* e *Normal-like*, ambos apresentaram incidência cumulativa intermediária, quando comparados aos outros subtipos, com 28,95% (IC 95%: 22,28–35,93) e 38,21% (IC 95%: 29,46–46,90), ao final dos 120 meses.

Observou-se, de maneira geral, maior amplitude nos intervalos de confiança à medida que o seguimento completo foi dividido em precoce e tardio, particularmente nos subtipos *Basal-like*, *Claudin-low* e *Normal-like*. Esse resultado decorre da diminuição do número de eventos e de indivíduos em risco ao final do período (Tabela [6](#) e [B4](#)). Tanto o modelo de riscos competitivos quanto as estimativas de incidência cumulativa de recorrência apresentam precisão adequada, com intervalos de confiança moderados no seguimento completo. No entanto, ao analisar os

seguimentos precoce e tardio, recomenda-se cautela, sobretudo nos subtipos com poucos eventos, que devem ser interpretados como exploratórios devido à maior incerteza estatística.

A instabilidade das estimativas tardias no subtipo *Basal-like* refletiu a presença simultânea de poucos eventos tardios ($n = 4$) e de óbitos sem recorrência ($n = 18$), que atuaram como eventos competidores (Tabela B4). Embora a direção do efeito no seguimento tardio (61–120 meses) tenha sido consistente entre o modelo do Cox causa-específico e o modelo de Fine-Gray (Tabela B2) (Cox HR: 0,266; IC 95%: 0,126–0,564; $p < 0,001$; Fine-Gray sHR: 0,264; IC 95%: 0,121–0,577; $p < 0,001$), a magnitude e a precisão das estimativas ainda permanecem suscetíveis a artefatos decorrentes da redução da amostra.

4.5 Avaliação do tempo de seguimento para as análises de sobrevida em 5 anos

Devido às diferenças estruturais entre os tipos de registros analisados, foi conduzida uma análise inicial, com o objetivo de avaliar o tempo de seguimento e definir qual a fonte de dados mais adequada para as análises de sobrevida.

Para descrever o tempo de seguimento do estudo, foram calculados os tempos individuais (pessoa-tempo), obtidos pela soma dos períodos (VAN DER VEN *et al.*, 2025). A distribuição desses tempos foi expressa por meio de medianas e intervalos interquartílicos. A distribuição de eventos também foi avaliada e expressa em contagem e porcentagem. O período observado incluiu os diagnósticos ocorridos entre 2000 e 2015, com acompanhamento até a data do encerramento informada pela fonte de dados, ocorrido em 13/06/2024 para o RHC Estadual (FUNDAÇÃO ONCOCENTRO DE SÃO PAULO, 2024) e 31/12/2021 para o RCBP Internacional (SEER PROGRAM, 2023). Destaca-se que para o RHC Nacional, nos casos em que o óbito foi registrado, a data do óbito foi considerada como término do seguimento. Para os demais registros, nos quais não havia informação de óbito, foi aplicada censura administrativa, utilizando-se a data da última atualização da base de dados (07/04/2025) para indicar a data de encerramento (INSTITUTO NACIONAL DE CÂNCER, 2026). Essa abordagem é metodologicamente consistente com a literatura (PUTTER; FIOCCO; GESKUS, 2007) e permitiu estimar o tempo potencial de observação no RHC Nacional, mesmo na ausência de registro de seguimento longitudinal sistemático, conforme demonstrado na Tabela 7.

Tabela 7 – Distribuição dos eventos e características do tempo de seguimento para diagnósticos realizados entre 2000 e 2015

Tipo de registro	Óbitos n (%)	Censuras n (%)	Tempo de seguimento (meses) ^(a)
			Mediana (IIQ)
RHC Estadual	40.297 (43,87%)	51.553 (56,13%)	80 (38–126)
RHC Nacional	68.130 (22,87%)	229.771 (77,13%)	158 (118–204)
RCBP Internacional	271.385 (36,35%)	475.182 (63,65%)	120 (80–174)

Notas: IIQ: Intervalo interquartil (Q1–Q3).

^(a)Corresponde ao tempo decorrido entre o diagnóstico e o último evento registrado (óbito ou data de censura). Nos casos em que o tempo de seguimento já constava de forma explícita na fonte original, esta informação foi mantida.

Fonte: Dados da pesquisa (2026).

Observou-se uma variação considerável na proporção de óbitos entre os tipos de registros analisados, com o **RHC** Estadual apresentando maior proporção de óbitos (43,87%), seguido pelo **RCBP** Internacional (36,35%), enquanto o **RHC** Nacional registrou a menor proporção (22,87%). Esse achado no **RHC** Nacional sugere que os dados disponíveis para acesso público nessa base de dados carecem de atualizações sistemáticas dos desfechos.

Quanto ao tempo de seguimento, os resultados demonstraram que tanto o **RHC** Estadual quanto o **RCBP** Internacional superaram o horizonte temporal de 60 meses necessários para análise de sobrevida em 5 anos (**RHC** Estadual: mediana = 80 meses; IIQ: 38–126; **RCBP** Internacional: mediana = 120 meses; IIQ: 80–174). Em contraste, no **RHC** Nacional, o tempo de seguimento estimado (mediana = 158; IIQ: 118–204) deve ser interpretado como exploratório. Considerando que o elevado número de censuras em data fixa (77,13%) e a baixa proporção de eventos nessa fonte de dados podem comprometer a precisão das inferências e estimativas de risco nos modelos multivariados, apenas os registros provenientes do **RHC** Estadual e **RCBP** Internacional foram utilizados nas análises de sobrevivência.

4.6 Caracterização das coortes: hospitalar e populacional

Para cada registro (**RHC** e **RCBP**), foram extraídas informações sobre idade, sexo, ano de diagnóstico, tempo de seguimento, estadiamento tumoral e indicadores de desfecho (óbitos por câncer, óbitos por outras causas e censuras). A escolha das variáveis foi orientada por meio de **DAG** (Figura C1) e buscou priorizar modelos de sobrevida parcimoniosos e clinicamente interpretáveis, considerando que a disponibilidade, a forma de coleta e a codificação dos dados variam entre os tipos de registro. Foram calculadas as estatísticas descritivas da amostra, nas

quais as variáveis categóricas foram apresentadas como contagem e porcentagem, e as variáveis contínuas como mediana e intervalo interquartilico.

O **RHC** Estadual incluiu 91.850 registros, com predomínio do sexo feminino (99,41%) e mediana de idade de 55 anos (IIQ: 46–66). Observou-se nessa coorte maior proporção de registros de tumores em estágio II (38,93%). Quanto aos desfechos, os dados apontaram para maior ocorrência de óbitos atribuídos ao câncer de mama (27,74%) do que por outras causas (16,13%).

Por outro lado, o **RCBP** Internacional incluiu 746.567 registros, com predomínio do sexo feminino (99,25%) e mediana de idade de 60 anos (IIQ: 50–71). Nessa coorte, verificou-se maior proporção de tumores em estágio I (48,21%). Destaca-se, ainda, o predomínio de óbitos por outras causas (22,14%) em relação aos óbitos atribuídos ao câncer de mama (14,21%). Os resultados detalhados dessa caracterização são apresentados na Tabela 8.

Tabela 8 – Caracterização demográfica e clínica das coortes: RHC Estadual e RCBP Internacional (Período: 2000–2015)

Variável	RHC Estadual	RCBP Internacional
	<i>n (%)</i>	<i>n (%)</i>
Sexo		
Feminino	91.311 (99,41%)	740.991 (99,25%)
Masculino	539 (0,59%)	5.576 (0,75%)
Idade (anos)		
Mediana (IIQ)	55 (46–66)	60 (50–71)
Estadiamento tumoral		
I	20.949 (22,81%)	359.895 (48,21%)
II	35.756 (38,93%)	257.989 (34,56%)
III	25.296 (27,54%)	88.792 (11,89%)
IV	8.191 (8,92%)	12.944 (1,73%)
Ausentes	1.658 (1,80%)	26.947 (3,61%)
Desfecho		
Censuras	51.553 (56,13%)	475.182 (63,65%)
Óbitos por câncer	25.481 (27,74%)	106.089 (14,21%)
Óbitos por outras causas	14.816 (16,13%)	165.296 (22,14%)

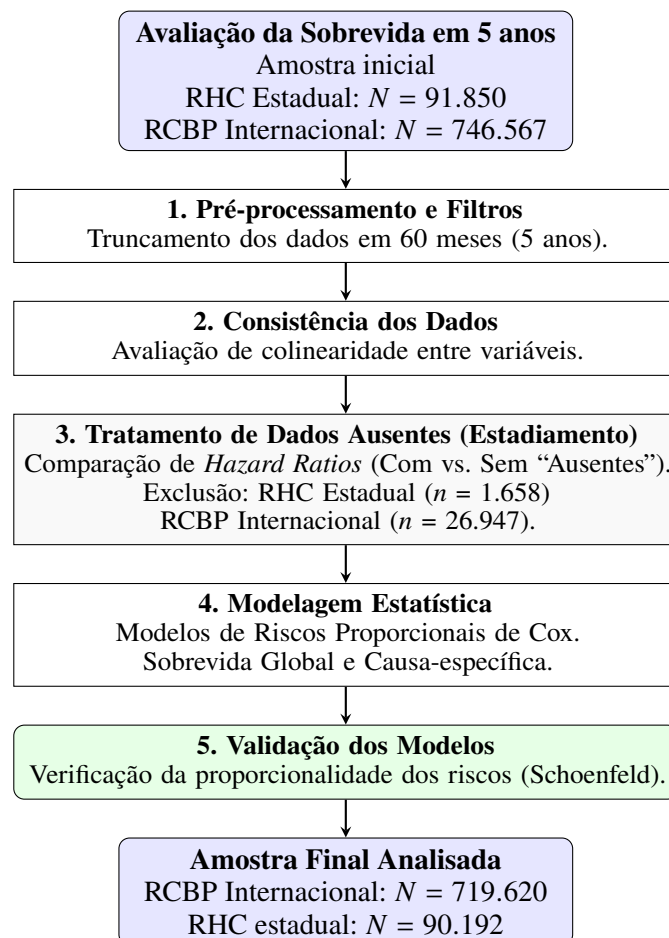
Notas: IIQ: Intervalo Interquartilico (Q1–Q3); RHC Estadual: $N=91.850$; RCBP Internacional: $N=746.567$. O período avaliado abrangeu os diagnósticos de câncer de mama ocorridos entre 2000 e 2015.

Fonte: Dados da pesquisa (2026).

4.7 Análise comparativa dos fatores prognósticos entre registros de câncer

Para comparar os fatores associados à sobrevida entre registros hospitalares e populacionais, os dados passaram por etapas de preparação analítica, que incluíram o truncamento dos dados em 5 anos após o diagnóstico, para estimar a sobrevida global e específica. A multicolinearidade entre as variáveis explicativas foi avaliada por meio do **GVIF** ajustado, cujo valor máximo foi de 1,003 (Tabela C1). Para as análises de sensibilidade a dados ausentes, foram consideradas duas abordagens: exclusão de observações com dados ausentes no estadiamento do tumor e inclusão de “Ausentes” como categoria explícita. Os resultados indicaram que a exclusão de dados com estadiamento ausente resultou em estimativas de risco semelhantes às obtidas na análise com a categoria explícita (Tabelas C2 e C3). O fluxo de preparação dos dados é apresentado na Figura 7, e os detalhes adicionais encontram-se no Apêndice C.

Figura 7 – Fluxo de preparação dos dados para análise comparativa entre registros.



Nota: N indica o tamanho da amostra total e n , o número de registros excluídos.

Fonte: Dados da pesquisa (2026).

As análises de sobrevida foram conduzidas por meio do modelo de riscos proporcionais de Cox. A proporcionalidade dos riscos foi avaliada por meio do teste de resíduos de Schoenfeld (Figuras C2, C3, C4 e C5). Os resultados indicaram que a hipótese de proporcionalidade dos riscos foi rejeitada para as variáveis idade e estadiamento no RHC Estadual (Tabelas C4 e C5). Para o RCBP Internacional, a proporcionalidade dos riscos foi rejeitada para todas as variáveis incluídas no modelo (idade, sexo, estadiamento do tumor e ano do diagnóstico). Nesse cenário, foram reportadas estimativas médias de efeito, indicadas como $HR_{\text{médio}}$. Os resultados das análises de sobrevida global em 5 anos são apresentados nas Tabelas 9 e 10.

Observou-se que a idade esteve consistentemente associada a maior risco de óbito em ambos os registros (hospitalar e populacional), com efeito mais pronunciado no RCBP Internacional em comparação com o RHC Estadual após ajuste para o estadiamento, sexo e ano de diagnóstico. Na coorte RHC Estadual (Tabela 9), o incremento médio do risco de morte para cada ano adicional de idade foi de 1,46% ($HR_{\text{médio}}$: 1,0146; IC 95%: 1,0136–1,0155; $p < 0,001$). Em contrapartida, na coorte RCBP Internacional (Tabela 10), o incremento foi de 5,03% no risco médio para cada ano ($HR_{\text{médio}}$: 1,0503; IC 95%: 1,0498–1,0508; $p < 0,001$).

Nos modelos univariados, o sexo masculino esteve associado a maior risco de óbito em ambos os registros. No RHC Estadual, esse efeito foi significativo na análise univariada ($HR_{\text{médio}}$: 1,336; IC 95%: 1,150–1,551; $p < 0,001$), mas perdeu o efeito após ajuste ($HR_{\text{médio}}$: 1,061; IC 95%: 0,914–1,233; $p = 0,436$). Já no RCBP Internacional, o sexo masculino manteve associação mesmo após ajuste ($HR_{\text{médio}}$: 1,287; IC 95%: 1,222–1,356; $p < 0,001$), indicando que, nesse registro, homens apresentam risco médio 28,7% maior de morte por todas as causas em 5 anos, quando comparados às mulheres (Tabela 10).

O estadiamento tumoral foi o fator mais fortemente associado à mortalidade global em ambos os contextos. No RHC Estadual, tumores diagnosticados em estágio IV tiveram uma associação ao risco de morte cerca de 19 vezes maior quando comparados ao estágio I ($HR_{\text{médio}}$: 19,423; IC 95%: 18,384–20,521; $p < 0,001$). De forma semelhante, no RCBP Internacional, o estágio IV elevou cerca de 15 vezes o risco médio de morte por todas as causas em 5 anos, quando comparado ao estágio I ($HR_{\text{médio}}$: 15,021; IC 95%: 14,661–15,390; $p < 0,001$).

O ano do diagnóstico esteve associado à redução no risco de morte no período estudado. No RHC Estadual, cada ano adicional esteve associado a uma redução de 2,38% no risco de morte por todas as causas ($HR_{\text{médio}}$: 0,9762; IC 95%: 0,9735–0,9789; $p < 0,001$). No RCBP Internacional, a redução foi de 2,67% ($HR_{\text{médio}}$: 0,9733; IC 95%: 0,9721–0,9746; $p < 0,001$). Esses achados sugerem melhora progressiva da sobrevida global ao longo do período estudado (Tabelas 9 e 10).

Tabela 9 – Análise de sobrevida global em 5 anos (Cox): fatores associados ao óbito no RHC Estadual.

Variável	Análise Univariada		Análise Multivariada	
	HR ^(a) (IC95%)	<i>p</i>	HR ^(a) (IC95%)	<i>p</i>
Idade (anos)^(b)	1,0129 (1,0120–1,0139)	<0,001	1,0146 (1,0136–1,0155)	<0,001
Sexo				
Feminino (Referência)				
Masculino	1,336 (1,150–1,551)	<0,001	1,061 (0,914–1,233)	0,436
Estadiamento tumoral				
I (Referência)				
II	2,184 (2,067–2,308)	<0,001	2,199 (2,081–2,324)	<0,001
III	6,213 (5,895–6,549)	<0,001	6,399 (6,071–6,746)	<0,001
IV	19,093 (18,071–20,173)	<0,001	19,423 (18,384–20,521)	<0,001
Ano do Diagnóstico^(b)	0,9795 (0,9768–0,9822)	<0,001	0,9762 (0,9735–0,9789)	<0,001

Nota: HR: Razão de Riscos (*Hazard Ratio*); IC95%: Intervalo de Confiança de 95%. Amostra total (*N*): 90.192; Número de eventos (óbitos em 5 anos): 23.752. Os diagnósticos ocorreram entre 2000 e 2015. As estimativas das variáveis contínuas foram apresentadas com quatro casas decimais para preservar a precisão do arredondamento.

^(a)Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade.

^(b)Para variáveis contínuas, o HR representa o risco associado ao incremento de uma unidade (1 ano).

Fonte: Dados da pesquisa (2026).

Tabela 10 – Análise de sobrevida global em 5 anos (Cox): Fatores associados ao óbito no RCBP Internacional

Variável	Análise Univariada		Análise Multivariada	
	HR ^(a) (IC95%)	<i>p</i>	HR ^(a) (IC95%)	<i>p</i>
Idade (anos)^(b)	1,0454 (1,0449–1,0459)	<0,001	1,0503 (1,0498–1,0508)	<0,001
Sexo				
Feminino (Referência)				
Masculino	2,044 (1,941–2,153)	<0,001	1,287 (1,222–1,356)	<0,001
Estadiamento tumoral				
I (Referência)				
II	1,778 (1,751–1,805)	<0,001	2,005 (1,975–2,035)	<0,001
III	4,208 (4,140–4,277)	<0,001	5,132 (5,049–5,217)	<0,001
IV	12,760 (12,455–13,073)	<0,001	15,021 (14,661–15,390)	<0,001
Ano do Diagnóstico^(b)	0,9701 (0,9688–0,9713)	<0,001	0,9733 (0,9721–0,9746)	<0,001

Nota: HR: Razão de Riscos (*Hazard Ratio*); IC95%: Intervalo de Confiança de 95%. Amostra total (*N*): 719.620; Número de eventos (óbitos em 5 anos): 103.991. Os diagnósticos ocorreram entre 2000 e 2015. As estimativas das variáveis contínuas foram apresentadas com quatro casas decimais para preservar a precisão do arredondamento.

^(a) Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade.

^(b) Para variáveis contínuas, o HR representa o risco associado ao incremento de uma unidade (1 ano).

Fonte: Dados da pesquisa (2026).

A análise da sobrevida específica por câncer em 5 anos revelou diferenças importantes entre as duas coortes. No **RHC** Estadual, a idade apresentou um incremento de 0,37% no risco de morte por câncer a cada ano adicional de idade, após ajuste por estadiamento, sexo e ano de diagnóstico (HR_{médio}: 1,0037; IC 95%: 1,0026–1,0048; *p* < 0,001. O sexo masculino perdeu efeito tanto na análise univariada (*p* = 0,084) quanto na multivariada (*p* = 0,879), sugerindo que não é um fator independente de risco nesse registro. Já o estadiamento do tumor apresentou magnitude muito maior na mortalidade específica: enquanto na análise global o estágio IV aumentou o risco médio em cerca de 19 vezes (Tabela 9), na análise específica, o modelo multivariado indicou que o risco aumentou para 43 vezes (HR_{médio}: 43,270; IC 95%: 39,849–46,985; *p* < 0,001). O ano do diagnóstico esteve associado à redução do risco médio de forma consistente em ambas as análises (redução de 2,81% no risco a cada ano adicional; HR_{médio}: 0,9719; IC 95%: 0,9688–0,9750; *p* < 0,001).

De forma semelhante, no **RCBP** Internacional, a idade também apresentou um efeito menor sobre a mortalidade específica. O incremento foi de 1,86% para cada ano a mais sobre a idade (HR_{médio}: 1,0186; IC 95%: 1,0179–1,0192; *p* < 0,001). O sexo masculino perdeu o efeito como variável independente na sobrevida específica após ajuste (*p* = 0,514). O risco médio associado a tumores diagnosticados em estágio IV foi cerca de 49 vezes maior quando comparado

ao estágio I na análise específica ($HR_{\text{médio}}$: 49,173; IC 95% 47,620–50,776; $p < 0,001$). O ano do diagnóstico manteve associação protetora, com redução de 2,96% a cada ano adicional ($HR_{\text{médio}}$: 0,9704; IC 95%: 0,9686–0,9721; $p < 0,001$). Mais detalhes são apresentados nas Tabelas 11 e 12.

Tabela 11 – Análise de sobrevida específica por câncer em 5 anos (Cox causa-específico): Fatores associados ao óbito no RHC Estadual

Variável	Análise Univariada		Análise Multivariada	
	HR ^(a) (IC95%)	<i>p</i>	HR ^(a) (IC95%)	<i>p</i>
Idade (anos)^(b)	1,0011 (1,0000–1,0022)	0,050	1,0037 (1,0026–1,0048)	<0,001
Sexo				
Feminino (Referência)				
Masculino	1,177 (0,979–1,417)	0,084	0,986 (0,819–1,186)	0,879
Estadiamento tumoral				
I (Referência)				
II	3,449 (3,170–3,753)	<0,001	3,398 (3,123–3,698)	<0,001
III	12,015 (11,080–13,030)	<0,001	12,010 (11,074–13,025)	<0,001
IV	42,990 (39,591–46,680)	<0,001	43,270 (39,849–46,985)	<0,001
Ano do Diagnóstico^(b)	0,9764 (0,9733–0,9796)	<0,001	0,9719 (0,9688–0,9750)	<0,001

Nota: HR: Razão de Riscos (*Hazard Ratio*); IC95%: Intervalo de Confiança de 95%. Amostra total (*N*): 90.192; Número de eventos (óbitos por câncer em 5 anos): 17.559. Os diagnósticos ocorreram entre 2000 e 2015. As estimativas das variáveis contínuas foram apresentadas com quatro casas decimais para preservar a precisão do arredondamento.

^(a) Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade.

^(b) Para variáveis contínuas, o HR representa o risco associado ao incremento de uma unidade (1 ano).

Fonte: Dados da pesquisa (2026).

Tabela 12 – Análise de sobrevida específica por câncer em 5 anos (Cox causa-específico): Fatores associados ao óbito no RCBP Internacional

Variável	Análise Univariada		Análise Multivariada	
	HR ^(a) (IC95%)	<i>p</i>	HR ^(a) (IC95%)	<i>p</i>
Idade (anos)^(b)	1,0087 (1,0081–1,0094)	<0,001	1,0186 (1,0179–1,0192)	<0,001
Sexo				
Feminino (Referência)				
Masculino	1,480 (1,362–1,608)	<0,001	0,973 (0,895–1,057)	0,514
Estadiamento tumoral				
I (Referência)				
II	3,687 (3,590–3,787)	<0,001	3,896 (3,793–4,001)	<0,001
III	12,970 (12,633–13,316)	<0,001	14,013 (13,647–14,389)	<0,001
IV	45,988 (44,540–47,483)	<0,001	49,173 (47,620–50,776)	<0,001
Ano do Diagnóstico^(b)	0,9651 (0,9634–0,9669)	<0,001	0,9704 (0,9686–0,9721)	<0,001

Nota: HR: Razão de Riscos (*Hazard Ratio*); IC95%: Intervalo de Confiança de 95%. Amostra total (*N*): 719.620; Número de eventos (óbitos por câncer em 5 anos): 55.201. Os diagnósticos ocorreram entre 2000 e 2015. As estimativas das variáveis contínuas foram apresentadas com quatro casas decimais para preservar a precisão do arredondamento.

^(a) Estimativas médias de efeito (*Average Hazard Ratios*) devido à violação da proporcionalidade.

^(b) Para variáveis contínuas, o HR representa o risco associado ao incremento de uma unidade (1 ano). **Fonte:** Dados da pesquisa (2026).

4.8 Comparação das probabilidades de sobrevida global e específica em 5 anos

Para complementar as análises de risco, foram estimadas as probabilidades de sobrevida aos 5 anos após o diagnóstico. Enquanto os modelos de Cox apresentados anteriormente quantificaram o risco médio de óbito associados aos fatores prognósticos, as curvas e probabilidades apresentadas nesta seção ilustram como esses riscos se traduzem em diferenças nas probabilidades de sobrevida ao longo dos 5 primeiros anos de acompanhamento.

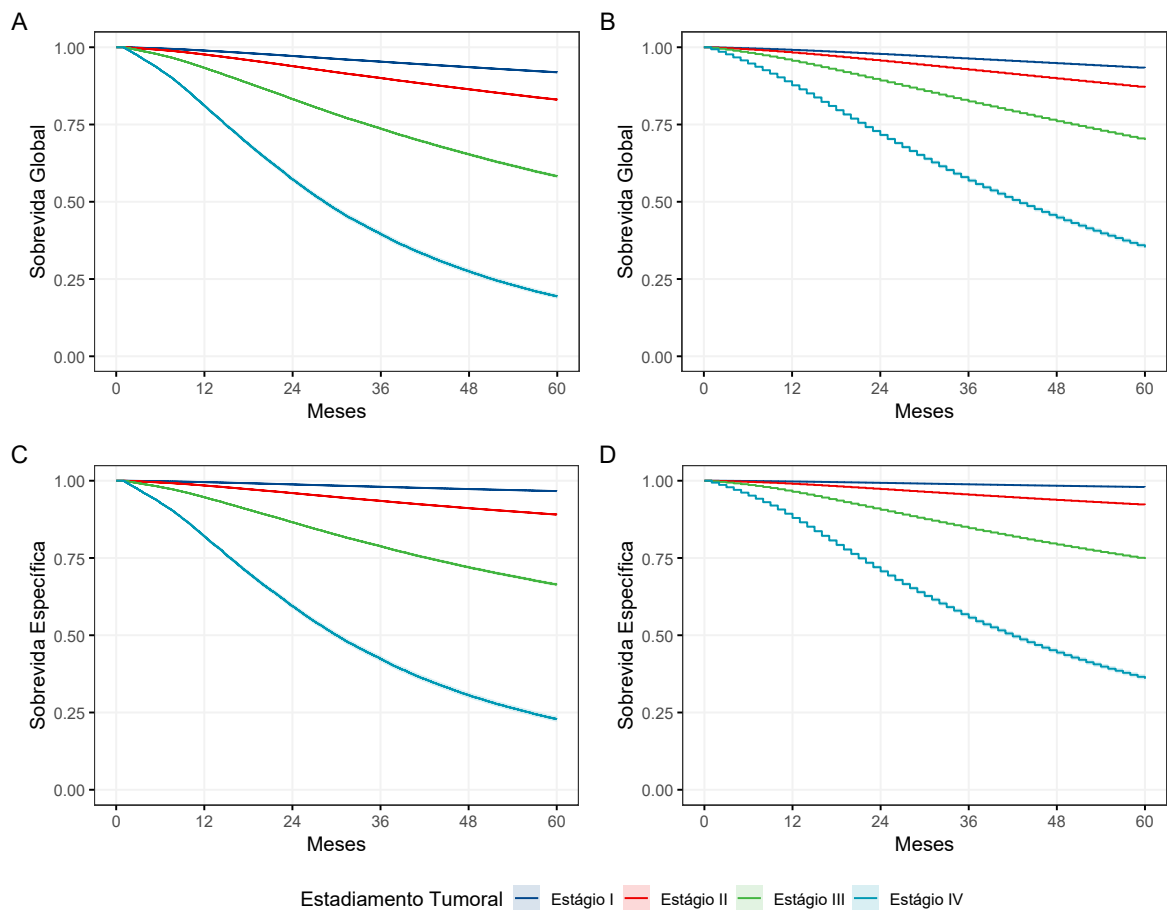
As curvas foram estimadas a partir dos modelos multivariados de sobrevida global e específica, em ambos os contextos de registro (hospitalar e populacional), e apresentam o gradiente prognóstico do estadiamento tumoral, no qual o avanço da doença associa-se à redução da probabilidade de sobrevida ao longo do tempo observado.

Conforme detalhado na Figura 8 e Tabela 13, as probabilidades de sobrevida específica foram superiores às probabilidades de sobrevida global em ambos os registros, refletindo o impacto dos óbitos por outras causas nas estimativas. No contexto do RHC Estadual, a sobrevida global variou de 91,90% (IC 95%: 91,50–92,30) no estágio I para 19,40% (IC 95%: 18,60–20,20)

no estágio IV (Painel A da Figura 8 e Tabela 13). Para a sobrevida específica, os valores foram ligeiramente superiores, variando de 96,60% (IC 95%: 96,40–96,90) no estágio I para 22,90% (IC 95%: 22,00–23,80) no estágio IV (Painel C da Figura 8 e Tabela 13).

No RCBP Internacional, os resultados mostraram um padrão semelhante: a sobrevida global variou de 93,30% (IC 95%: 93,20–93,40) no estágio I para 35,20% (IC 95%: 34,40–36,00) no estágio IV (Painel B da Figura 8 e Tabela 13). Já a sobrevida específica variou de 97,90% (IC 95%: 97,90–98,00) no estágio I a 35,90% (IC 95%: 35,10–36,80) no estágio IV (Painel D da Figura 8 e Tabela 13).

Figura 8 – Curvas ajustadas de sobrevida global e específica em 5 anos, segundo estadiamento tumoral



Nota: As curvas representam a probabilidade de sobrevida estimada pelo modelo de riscos proporcionais de Cox, ajustadas por idade, sexo e ano de diagnóstico: (A) Sobrevida Global - RHC Estadual; (B) Sobrevida Global - RCBP Internacional; (C) Sobrevida Específica - RHC Estadual; (D) Sobrevida Específica - RCBP Internacional. As linhas representam o estadiamento do tumor, estratificado pelos estágios I ao IV.

Fonte: Dados da pesquisa (2026).

Tabela 13 – Probabilidade de sobrevida em 5 anos segundo estadiamento tumoral

Fonte / Estágio	Sobrevida Global (%) (IC 95%)	Sobrevida Específica (%) (IC 95%)
RHC Estadual		
Estágio I	91,90 (91,50–92,30)	96,60 (96,40–96,90)
Estágio II	83,10 (82,70–83,50)	89,10 (88,70–89,40)
Estágio III	58,20 (57,60–58,90)	66,40 (65,80–67,00)
Estágio IV	19,40 (18,60–20,20)	22,90 (22,00–23,80)
RCBP Internacional		
Estágio I	93,30 (93,20–93,40)	97,90 (97,90–98,00)
Estágio II	87,00 (86,90–87,10)	92,20 (92,10–92,30)
Estágio III	70,00 (69,70–70,30)	74,70 (74,40–75,00)
Estágio IV	35,20 (34,40–36,00)	35,90 (35,10–36,80)

Nota: Valores expressos em porcentagem e intervalo de confiança de 95%, ajustados por idade, sexo e ano de diagnóstico.

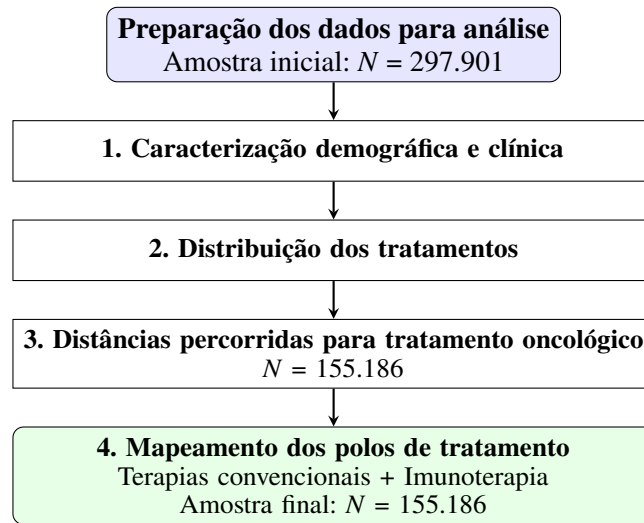
Fonte: Dados da pesquisa (2026).

4.9 Caracterização demográfica, clínico-patológica e terapêutica dos Registros Hospitalares de Câncer do Brasil

Originalmente, os registros do [RHC](#) Nacional seriam utilizados em análises de sobrevida. Contudo, as análises preliminares de consistência revelaram limitações relacionadas à atualização dos desfechos, inviabilizando estimativas confiáveis de risco. Dessa forma, essa seção concentra-se na caracterização dos registros hospitalares de câncer de mama na rede oncológica nacional, consolidando informações de 297.901 registros. Manteve-se a estrutura amostral originalmente estabelecida, incluindo os diagnósticos ocorridos entre 01/01/2000 e 31/12/2015, preservando a homogeneidade metodológica em relação aos demais registros de câncer avaliados neste trabalho.

As análises realizadas com os registros do [RHC](#) Nacional incluíram a caracterização demográfica e clínico-patológica da coorte, a descrição das modalidades terapêuticas, a avaliação dos deslocamentos para tratamento e o mapeamento da distribuição espacial dos serviços oncológicos. O fluxo das etapas analíticas é apresentado na Figura 9.

Figura 9 – Fluxo análise dos registros de câncer de mama no RHC Nacional.



Nota: N indica o tamanho da amostra em cada etapa.

Fonte: Dados da pesquisa (2026).

4.9.1 Perfil demográfico e clínico

A análise dos registros no **RHC** Nacional revelou uma maior proporção de diagnósticos de câncer de mama na faixa etária ≤ 49 anos, representando 36,15% dos registros avaliados, enquanto o diagnóstico em idades mais avançadas, ≥ 70 anos, ocorreu em 16,76% das observações. Houve amplo predomínio do sexo feminino, correspondendo a 98,96% dos registros, enquanto a população masculina representou 1,04% da amostra.

Em relação ao perfil clínico, os dados de estadiamento tumoral indicaram maior proporção de tumores classificados como estágio II (33,20%) e III (25,43%), respectivamente. Ausência de estadiamento tumoral foi registrada em 17,26% das observações.

A análise da distribuição geográfica demonstrou que 51,02% dos registros de câncer de mama no **RHC** Nacional foram oriundos da região Sudeste, seguida pela região Nordeste, com 23,69% e Sul, com 18,42%. As regiões Norte e Centro-Oeste apresentaram menores proporções de registros, com 3,79% e 3,08%, respectivamente. Por fim, a avaliação do deslocamento revelou que 52,09% da amostra foi tratada fora da cidade de domicílio. Os detalhes da análise são apresentados na Tabela 14.

Tabela 14 – Caracterização demográfica e clínica do RHC Nacional (Período: 2000 a 2015)

Variável	<i>n</i>	%
Faixa Etária (anos)		
≤49	107.688	36,15
50–59	79.439	26,67
60–69	60.841	20,42
≥70	49.933	16,76
Sexo		
Feminino	294.809	98,96
Masculino	3.092	1,04
Estadiamento tumoral		
I	49.417	16,59
II	98.910	33,20
III	75.762	25,43
IV	22.404	7,52
Ausente	51.408	17,26
Região de Residência		
Centro-Oeste	9.187	3,08
Nordeste	70.560	23,69
Norte	11.298	3,79
Sudeste	151.988	51,02
Sul	54.868	18,42
Tratamento na cidade onde mora		
Não	155.186	52,09
Sim	142.715	47,91

Nota: *N* = 297.901.

Fonte: Dados da pesquisa (2026).

4.9.2 Distribuição dos tratamentos realizados

A cirurgia foi a modalidade terapêutica mais utilizada (66,79%), principalmente na região Sudeste (75,66%), seguida pelas regiões Nordeste (62,32%) e Norte (57,67%). A quimioterapia apresentou um padrão semelhante de utilização, em que 66,36% dos registros avaliados indicaram esta modalidade como parte do tratamento, alcançando maior proporção de utilização nas regiões Norte (69,54%), Nordeste (69,08%) e Sudeste (68,67%).

A radioterapia foi utilizada em 48,23% dos registros avaliados (Tabela 15), com maior utilização nas regiões Norte (54,11%) e Nordeste (50,37%), seguidas pelas regiões Sudeste (49,13%), Centro-Oeste (49,77%) e Sul (42,08%) (Figura 10).

A hormonioterapia apresentou um padrão de utilização semelhante à radioterapia, sendo reportada em 43,12% dos registros, entretanto, destaca-se o contraste observado nas regiões Centro-Oeste e Norte, com menor utilização dessa modalidade (Região Centro-Oeste: 28,53%; Região Norte: 27,24%), quando comparada às regiões Sudeste (52,50%), Nordeste (36,80%) e Sul (31,47%).

A imunoterapia foi a modalidade com menor utilização no período observado (0,53%). O uso concentrou-se principalmente nas regiões Nordeste (0,79%) e Sudeste (0,58%), seguidas pelas regiões Sul (0,24%), Norte (0,21%) e Centro-Oeste (0,03%). A distribuição regional e percentual das modalidades terapêuticas utilizadas no período avaliado é demonstrada na Tabela 15 e Figura 10. Detalhes adicionais desta análise são apresentados na Tabela D1 do Apêndice D.

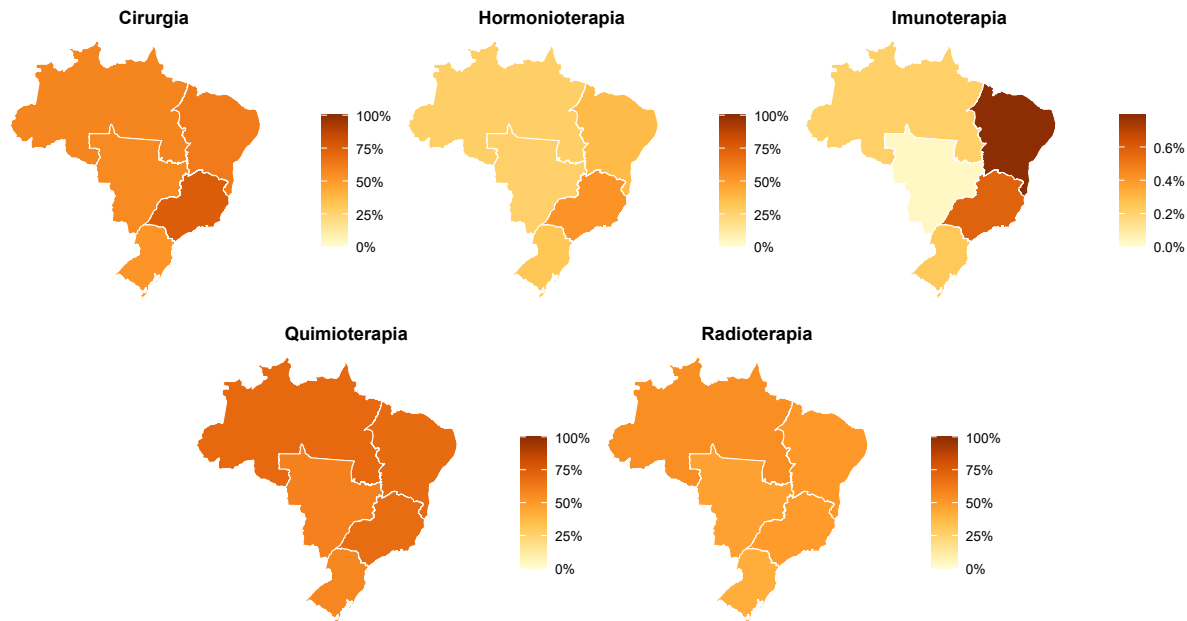
Tabela 15 – Frequência de utilização das modalidades terapêuticas no RHC Nacional (Período: 2000 a 2015)

Modalidade Terapêutica	Sim n (%)	Não n (%)	Ausente n (%)
Cirurgia	198.989 (66,79%)	98.364 (33,02%)	548 (0,18%)
Quimioterapia	197.699 (66,36%)	99.654 (33,45%)	548 (0,18%)
Radioterapia	143.666 (48,23%)	153.687 (51,59%)	548 (0,18%)
Hormonioterapia	128.448 (43,12%)	168.906 (56,70%)	547 (0,18%)
Imunoterapia	1.593 (0,53%)	295.760 (99,28%)	548 (0,18%)

Nota: Os percentuais baseiam-se no universo total de registros amostrados no RHC Nacional ($N = 297.901$). Os valores de cada linha são independentes, visto que os tratamentos não são mutuamente exclusivos, e um indivíduo pode ter recebido terapias combinadas.

Fonte: Dados da pesquisa (2026).

Figura 10 – Perfil regional de utilização das modalidades de tratamento para o câncer de mama no Brasil (Período: 2000 a 2015)



Nota: O gráfico apresenta a distribuição espacial da proporção de registros de indivíduos que realizaram cada modalidade terapêutica. A escala de cores utiliza um gradiente sequencial, no qual tonalidades mais claras representam menores proporções e tonalidades mais escuras representam maiores proporções de utilização da terapia. Para fins de comparabilidade direta, as escalas de legenda foram padronizadas de forma fixa (0% a 100%) para as modalidades de cirurgia, hormonioterapia, quimioterapia e radioterapia. Excepcionalmente, para viabilizar a visualização dos padrões geográficos da imunoterapia, cujas proporções observadas foram inferiores a 1%, esse painel utiliza uma escala de legenda própria, ajustada ao intervalo de valores observado.

Fonte: Dados da pesquisa (2026).

4.9.3 Distâncias percorridas para tratamento oncológico

Para estimar e descrever as distâncias de deslocamento para tratamento oncológico, foi calculada a distância entre o município de residência e o município de tratamento de cada registro. As coordenadas geográficas dos municípios brasileiros foram obtidas a partir do pacote `geobr` no RStudio (PEREIRA *et al.*, 2026), que disponibiliza dados oficiais do IBGE em formato espacial. A partir dessas coordenadas, foram extraídos os centróides municipais (latitude e longitude) e vinculados aos registros hospitalares por meio da padronização dos códigos de município (BRASIL, 2025). Devido a restrições éticas e de sigilo de dados, o endereço domiciliar não estava disponível na fonte de dados do RHC Nacional, consequentemente, para os registros em que o município de residência e o de tratamento eram coincidentes, os percursos não foram calculados. Diante dessa limitação, foram exploradas as trajetórias de 155.186 registros de câncer que indicaram deslocamento no tratamento oncológico.

É importante destacar que os registros hospitalares de câncer são organizados por tumor e não por paciente. Assim, embora cada tumor esteja necessariamente vinculado a um indivíduo, um mesmo paciente pode estar representado em mais de um registro, caso tenha apresentado mais de um tumor primário ao longo do período analisado. Consequentemente, as análises de deslocamento refletem o número de registros de tumores, cujos indivíduos demandaram tratamento fora do município de residência, e não o número absoluto de indivíduos distintos. Como os dados são públicos e anonimizados, não foi possível analisar as distâncias por pacientes. Ainda assim, cada registro corresponde a um evento real, e cada tumor exigiu um tratamento específico e uma necessidade de acesso aos serviços de saúde. Portanto, os resultados a seguir são representativos da carga de deslocamentos necessária ao tratamento dos tumores de mama.

O cálculo das distâncias rodoviárias foi realizado utilizando o pacote *osrm* no RStudio e permitiu estimar rotas reais de deslocamento com base nos dados do OpenStreetMap (GIRAUD *et al.*, 2025; CONTRIBUTORS, 2025). Este procedimento forneceu a distância percorrida em quilômetros (km). A categorização das distâncias foi fundamentada tanto na literatura (FONSECA *et al.*, 2022) quanto nas análises das medianas e dos intervalos interquartílicos da amostra.

Entre as modalidades terapêuticas avaliadas, a hormonioterapia foi a que apresentou menor distância percorrida, com mediana de 74,6 km (IIQ: 40,0–159,0). Em contrapartida, a imunoterapia foi a modalidade que demandou maior deslocamento, com mediana de 114,0 km (IIQ: 45,0–218,0), seguida pela radioterapia, com mediana de 88,2 km (IIQ: 45,0–178,0). Quimioterapia e cirurgia apresentaram medianas de deslocamentos de 84,0 e 81,8 km respectivamente, e os detalhes são apresentados na Tabela 16.

Tabela 16 – Distância percorrida (km) por indivíduos vinculados aos registros de câncer e tratados fora do município de residência (Período: 2000 a 2015)

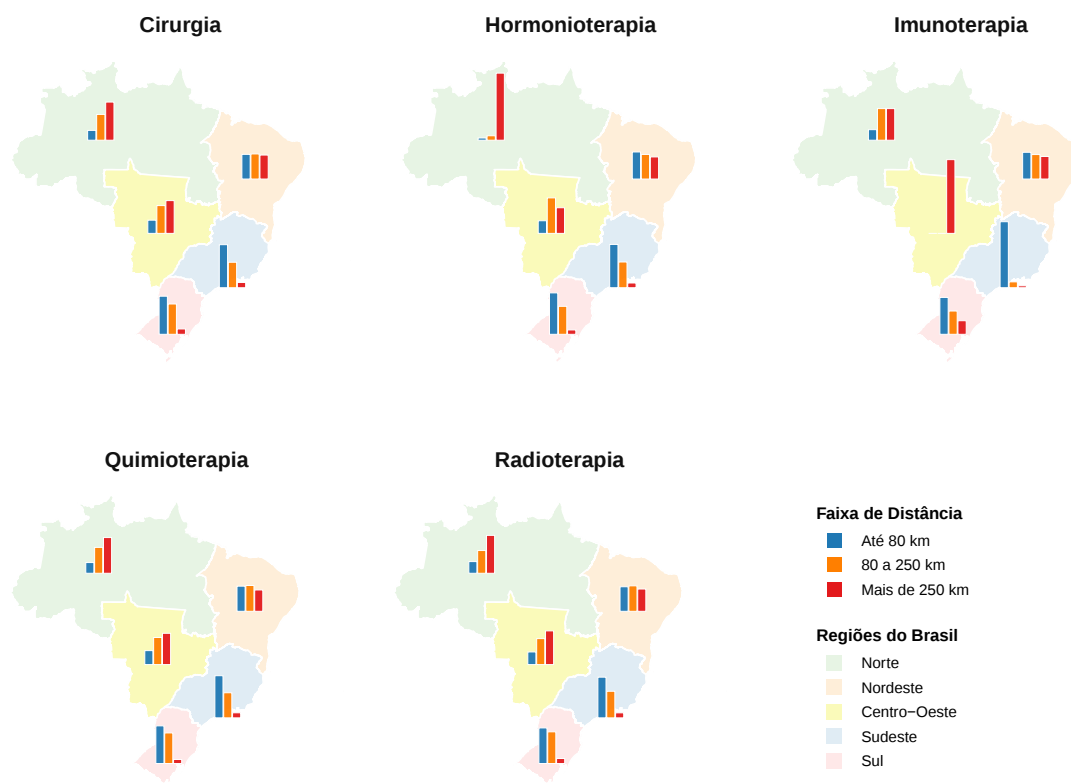
Modalidade Terapêutica	Distância Mediana (km)	IIQ (Q1–Q3)
Cirurgia	81,8	42,4–173,0
Quimioterapia	84,0	43,0–171,0
Radioterapia	88,2	45,0–178,0
Hormonioterapia	74,6	40,0–159,0
Imunoterapia	114,0	45,0–218,0

Nota: IIQ: Intervalo Interquartílico (Q1: 25%; Q3: 75%). A tabela foi construída com base em um total de 155.186 registros. As modalidades terapêuticas não são mutuamente exclusivas, pois um mesmo indivíduo pode ter realizado terapias combinadas.

Fonte: Dados da pesquisa (2026).

A análise da distribuição das distâncias percorridas por região brasileira evidenciou cenários distintos de acessibilidade ao tratamento oncológico no período observado. Os detalhes são apresentados na Figura 11.

Figura 11 – Distribuição das distâncias percorridas por indivíduos que realizaram tratamento oncológico fora de seu município de residência. (Período: 2000 a 2015)



Nota: A figura apresenta o perfil de deslocamento de indivíduos que buscaram tratamento oncológico fora de seu município de residência. Para cada região do Brasil e modalidade de tratamento, as barras verticais indicam o percentual de pacientes distribuídos proporcionalmente em três faixas de distância: até 80 km, de 80 a 250 km e acima de 250 km. A altura das barras reflete o peso de cada distância dentro da própria região, destacando as diferenças regionais na necessidade de deslocamento para tratamento oncológico.

Fonte: Dados da pesquisa (2026).

De maneira geral, a radioterapia foi a modalidade com maior deslocamento (acima de 250 km) nas regiões Norte (52,41%) e Centro-Oeste (46,62%). A imunoterapia, modalidade com menor proporção de utilização no período avaliado, concentrou deslocamentos de até 80 km nas regiões Sul (50,00%), Sudeste (45,97%) e Nordeste (36,24%). No entanto, essas proporções não necessariamente superaram as observadas nas demais modalidades terapêuticas, devido ao reduzido número absoluto de registros, o que reflete a distribuição dentro de sua própria categoria e não permite inferir maior acessibilidade em relação às demais terapias (Tabela D2).

A região Nordeste caracterizou-se por uma distribuição relativamente equilibrada entre as faixas de distância avaliadas. A cirurgia, por exemplo, dividiu-se entre os trajetos de até 80 km (33,44%), de 80 a 250 km (34,04%) e acima de 250 km (32,52%). Distribuição análoga foi

observada para as demais modalidades nessa região.

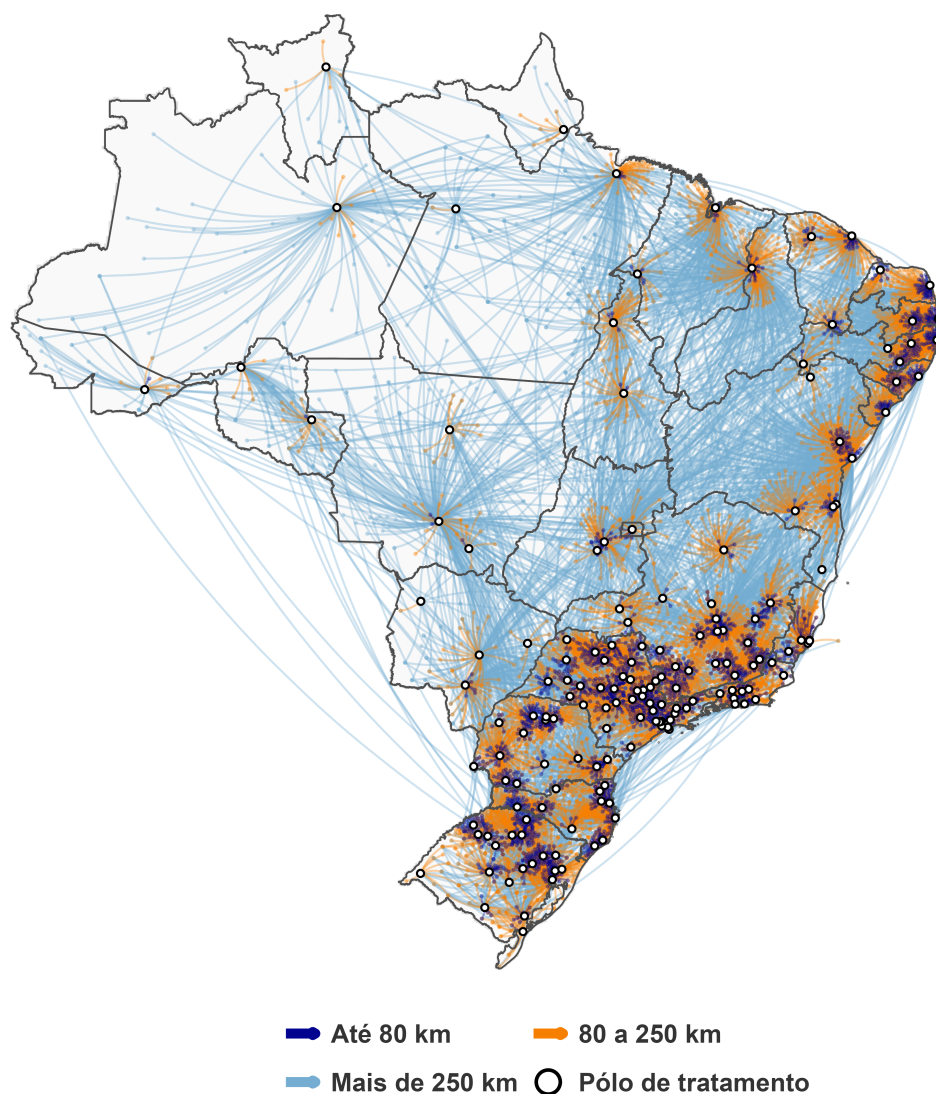
As regiões Sudeste e Sul exibiram predomínio de deslocamentos de curta distância. No Sudeste, a faixa de até 80 km compõe a maior parte das modalidades terapêuticas avaliadas, enquanto deslocamentos acima de 250 km foram residuais, com proporções máximas observadas em 13,13% dos registros para a modalidade imunoterapia. Os deslocamentos acima de 250 km também se mostraram reduzidos na região Sul, variando entre 5,61% dos registros para a modalidade quimioterapia e 18,42% para imunoterapia (Figura 11 e Tabela D2).

De forma complementar, a distribuição espacial dos polos de tratamento oncológico e dos fluxos de deslocamento foi mapeada e apresentada nas Figuras 12 e 13. O mapa elaborado confirma visualmente os padrões de longa distância observados nas regiões Norte e Centro-Oeste.

Embora, de maneira geral, as distâncias percorridas não tenham excedido a mediana de 114 km (Tabela 16), algumas linhas que cruzam as fronteiras estaduais indicam trajetos migratórios rumo a polos de tratamento em outros estados (mais de 250 km). Os fluxos de média distância (80 a 250 km) demonstram a conexão entre os municípios periféricos aos polos de tratamento. Observa-se, ainda, que os polos de oferta de terapias convencionais (cirurgia, quimioterapia, radioterapia e hormonioterapia) estão densamente agrupados nas regiões Sudeste, Sul e Nordeste.

O mapa apresentado na Figura 13 demonstrou a regionalização da imunoterapia no período observado. Os polos de tratamento concentraram-se nas regiões Sudeste, Sul e Nordeste, com escassez de oferta nas regiões Norte e Centro-Oeste. Os detalhes são apresentados a seguir.

Figura 12 – Mapeamento dos polos de tratamento e fluxos de deslocamento para terapias convencionais em pacientes com câncer de mama (Período: 2000 a 2015)



Nota: O mapa ilustra o fluxo de deslocamento para a realização de terapias convencionais (cirurgia, quimioterapia, radioterapia e hormonioterapia). Cada linha curva representa a trajetória do indivíduo, do seu município de residência (origem) até o município de atendimento (destino). Os pontos circulares brancos com borda preta demarcam a localização geográfica dos polos de saúde que concentram a oferta desses serviços. As distâncias percorridas estão divididas em três faixas de deslocamento, onde cada cor indica a amplitude do trajeto percorrido da origem ao destino.

Fonte: Dados da pesquisa (2026).

Figura 13 – Mapeamento dos polos de tratamento e fluxos de deslocamento para imunoterapia em pacientes com câncer de mama (Período: 2000 a 2015)



Nota: O mapa ilustra o fluxo de deslocamento para a realização de imunoterapia. Cada linha curva representa a trajetória do indivíduo, do seu município de residência (origem) até o município de atendimento (destino). Os pontos circulares brancos com borda preta demarcam a localização geográfica dos polos de saúde que concentram a oferta desse serviço específico. As distâncias percorridas estão divididas em três faixas de deslocamento, onde cada cor indica a amplitude do trajeto percorrido da origem ao destino.

Fonte: Dados da pesquisa (2026).

5 DISCUSSÃO

5.1 A robustez dos registros moleculares

A ausência de dados moleculares em larga escala frequentemente limita o alcance das análises epidemiológicas. Ao investigar uma fonte de dados genômicos, a inclusão dos subtipos moleculares intrínsecos identificados pelo teste [PAM50](#) revelou o impacto prognóstico dessas assinaturas genômicas sobre os desfechos de recorrência do câncer de mama.

A coorte molecular indicou a prevalência dos tumores luminais, com destaque para o Luminal A (Tabela 2). Esse subtipo, caracterizado por um perfil clínico mais indolente, foi compatível com os achados descritos na caracterização da coorte (Tabela 3), onde se observou ausência de comprometimento linfonodal na maioria dos registros avaliados, predomínio de graus histológicos baixos ou intermediários, idade mais avançada ao diagnóstico e status pós-menopáusico. Esse mesmo perfil clínico é descrito por outros autores ([CHAS *et al.*, 2018](#); [CAREY *et al.*, 2006](#)).

Destaca-se que o subtipo luminal A apresenta crescimento lento e menor invasividade na fase inicial da doença, contudo, o risco de recorrência, mesmo reduzido em comparação com o subtipo luminal B e com outros subtipos intrínsecos, pode se estender muito além do período adjuvante padrão de 5 anos ([GAO; SWAIN, 2018](#); [RUEDA *et al.*, 2019](#)).

Neste estudo, as recorrências do subtipo Luminal B concentraram-se nos primeiros 5 anos de seguimento, mas a curva de incidência cumulativa não se estabilizou, permanecendo ascendente e alcançando 44,52% aos 10 anos de seguimento (Figura 6). Estudos anteriores confirmam que tumores Luminal B apresentam recidivas precoces, mas podem continuar a acumular novos eventos em seguimentos prolongados ([RUEDA *et al.*, 2019](#); [LI *et al.*, 2016](#)). Já foi demonstrado que tumores Luminal B são mais agressivos que o Luminal A porque apresentam maior taxa proliferativa ([Ki-67](#) elevado), menor expressão de receptores de progesterona e, em certos casos, amplificação ou superexpressão de [HER2](#) ([CHEANG *et al.*, 2009](#); [LI *et al.*, 2016](#)). Esses fatores podem explicar o crescimento rápido e a maior agressividade biológica observados nesse subtipo.

A hormonioterapia foi a base do tratamento dos tumores luminais na coorte molecular (Tabela 3). Esse padrão é biologicamente compatível com a expressão de receptores hormonais observada nesses subtipos, os quais atualmente têm a terapia endócrina como principal estratégia terapêutica ([CUCCINIELLO *et al.*, 2022](#); [LOIBL *et al.*, 2024](#)).

Em contrapartida, a exploração desta coorte revelou o perfil mais agressivo de *Basal-like* e [HER2](#) enriquecido, com maior proporção de registros em mulheres mais jovens, tumores maiores, de alto grau histológico e comprometimento linfonodal extenso (Tabela 3). Esses achados são

consistentes com a descrição de estudos recentes (QIU *et al.*, 2024; COSAR *et al.*, 2022).

O maior uso de quimioterapia observado nos subtipos *Basal-like* e HER2 enriquecido reflete a ausência de alvos terapêuticos hormonais no primeiro e a superexpressão ou amplificação do HER2 no segundo (CARVALHO, 2023; HOHMANN *et al.*, 2025). Como esses subtipos intrínsecos são frequentemente associados a tumores maiores, de alto grau histológico e com maior comprometimento linfonodal, tornam-se necessários esquemas terapêuticos mais intensivos. Antes da introdução de terapias-alvo, a quimioterapia era indispensável para tumores HER2 enriquecido. Mesmo hoje, os protocolos recomendam a combinação de quimioterapia e terapia anti-HER2, devido ao elevado risco de progressão precoce (STANOWICKA-GRADA; SENKUS, 2023). Este achado alinha-se com o elevado risco médio de recorrência no seguimento precoce demonstrado nesta análise (Tabela 5).

É importante contextualizar os resultados no que diz respeito à descrição dos tratamentos: a coorte molecular, cujos dados foram extraídos do estudo METABRIC, reúne tumores de cinco centros do Reino Unido e do Canadá. Estima-se que os tumores da coorte original tenham sido coletados entre 1977 e 2005 (CURTIS *et al.*, 2012; MUKHERJEE *et al.*, 2018), período anterior à caracterização dos subtipos intrínsecos (PARKER *et al.*, 2009; PEROU *et al.*, 2000). Assim, a classificação molecular por PAM50 foi realizada retrospectivamente, utilizando material tumoral armazenado, e as características terapêuticas apresentadas nesta análise refletem as decisões tomadas com base nos critérios disponíveis na época.

O comportamento indolente do subtipo *Claudin-low* observado nesta coorte (Tabela 5), assemelhando-se ao Luminal A, encontra sustentação na literatura. Embora inicialmente descrito como um subtipo intrínseco de prognóstico desfavorável por apresentar características de transição epitélio-mesenquimal e células-tronco tumorais (PRAT *et al.*, 2010), investigações subsequentes demonstraram que o subtipo *Claudin-low* não deve ser tratado como um subtipo intrínseco único, pois apresenta fenótipos distintos, o que pode alterar o prognóstico (FOUGNER *et al.*, 2020; POMMIER *et al.*, 2020; VOUTSADAKIS, 2023). Do ponto de vista biológico, os tumores classificados como *Claudin-low* caracterizam-se por estabilidade genômica, baixos níveis de carga mutacional e baixas taxas de proliferação celular, além de um massivo infiltrado de células estromais e imunes. Este perfil biológico e a heterogeneidade interna, ou seja, a existência de subgrupos de tumores *Claudin-low*, podem explicar a ausência de diferença estatisticamente significativa em relação ao subtipo Luminal A encontrada no presente estudo.

Normal-like apresentou comportamento clínico e prognóstico intermediário em relação aos subtipos Luminal A e HER2 enriquecido, com incidência cumulativa de recorrência de 38,21% aos 10 anos (Tabela 6). Este subtipo é pouco explorado na literatura quanto à incidência cumulativa e à análise de risco. Inclusive, o estudo seminal de Parker *et al.* (2009) descrevia o subtipo *Normal-like* como provável artefato decorrente de contaminação da amostra com tecido mamário normal durante as análises de perfil de expressão gênica. Por outro lado, investigações mais recentes que exploraram a biologia tumoral desse subtipo revelam que o tumor

é caracterizado por elevada atividade estromal e imune (LI; WU; HAN, 2022).

Um achado relevante apontado nesta análise foi que os subtipos não luminais (*Basal-like* e *HER2* enriquecido) e Luminal B apresentaram maior risco médio de recorrência nos primeiros 60 meses, seguido de declínio ou perda de significância estatística no período tardio (Tabela 5). Esse comportamento encontra suporte no trabalho de Rueda *et al.* (2019), ao demonstrar que, uma vez superado o marco crítico de 5 anos livres da doença, o risco residual declina drasticamente em subtipos não luminais, ao passo que tumores luminais mantêm risco tardio persistente. Do ponto de vista clínico, compreender essa dinâmica temporal justifica a intensificação do monitoramento nos primeiros anos para pacientes com subtipos não luminais, enquanto, para tumores luminais, recomenda-se uma vigilância estendida e de longo prazo, dado o risco persistente de recorrência tardia (LOIBL *et al.*, 2024).

De forma similar ao observado nos subtipos moleculares, o grau histológico 3 não manteve significância estatística com a recorrência no seguimento tardio (Tabela 5). Achados análogos foram descritos por Yekedüz *et al.* (2022), que, embora tenham utilizado uma abordagem estatística distinta (*Odds Ratio*), documentaram associação do grau histológico intermediário e alto grau com a recorrência precoce (OR: 3,09; IC 95%: 1,87—5,12; $p < 0,001$), e ausência de significância no período tardio (OR: 0,66; IC 95%: 0,37—1,16; $p = 0,15$) (YEKEDÜZ *et al.*, 2022). Cabe salientar que *Odds Ratio* e estimativas derivadas de modelos de risco de subdistribuição (sHR) têm interpretações distintas, o que limita comparações diretas; ainda assim, a concordância temporal entre os estudos reforça a hipótese de maior impacto prognóstico do grau histológico no risco de recorrência precoce. Uma possível explicação biológica é que o grau histológico reflete características associadas à cinética de proliferação celular e à agressividade biológica inicial da doença (RAKHA; TOSS; QUINN, 2022; JAROENSRI *et al.*, 2022). Dessa forma, os tumores de grau 3 podem apresentar intensa atividade proliferativa nos primeiros 60 meses após o diagnóstico, o que aumenta o risco de recorrência.

Nesta análise, observou-se também uma associação marginal ($p = 0,059$) entre grau histológico 2 e o risco de recorrência no período tardio. Esse achado sugere um efeito mais distribuído ao longo do tempo observado, haja vista que o seguimento completo apresentou uma associação significativa entre o grau histológico 2 e o risco de recorrência (sHR_{médio}: 1,429; IC 95%: 1,039—1,966; $p = 0,028$). Embora terapias iniciais efetivamente reduzam a atividade mitótica tumoral (ZOMBORI *et al.*, 2025), a persistência de risco tardio de recorrência em alguns pacientes é apoiada no conceito de dormência de células tumorais residuais, descrita em revisão recente (DALLA *et al.*, 2023). De acordo com os autores, a dormência pode ser induzida após tratamentos direcionados ao tumor ou pelo próprio microambiente tumoral, permitindo que essas células reduzam drasticamente seu metabolismo e taxa mitótica, permanecendo em estado de latência por anos após a implementação do tratamento. A presença desses e outros mecanismos supressores do tumor pode justificar a ausência de significância estatística do grau histológico em seguimentos de longo prazo. Por outro lado, a interpretação desses achados também deve

considerar a redução no número de eventos no seguimento tardio, que diminui o poder estatístico (Tabela B4).

Apesar do protagonismo dos subtipos moleculares intrínsecos, os achados desta análise demonstraram que o tamanho do tumor e o comprometimento linfonodal mantiveram um papel preditor independente e relevante para o risco de recorrência. Tumores maiores refletem maior carga tumoral, aumentando a probabilidade de disseminação de células neoplásicas para linfonodos regionais e sítios distantes. Já o comprometimento linfonodal representa um marcador da extensão anatômica da doença e da capacidade metastática do tumor, justificando a associação dessas características com a probabilidade de recidiva. Essa interpretação encontra suporte na literatura: em estudo populacional que avaliou a recorrência por até 32 anos após o diagnóstico, Pedersen *et al.* (2022) identificaram a carga linfonodal elevada e o tamanho tumoral superior a 20 mm como preditores clínico-patológicos de recorrência tardia, concluindo que pacientes com tais características beneficiam-se de vigilância prolongada e estratégias terapêuticas mais intensivas.

Essa complementariedade entre variáveis anatômicas e moleculares reflete a própria evolução histórica dos sistemas de classificação da doença. Consolida-se, assim, o entendimento trazido pela 8ª edição do AJCC, que não descartou o sistema TNM tradicional, mas o refinou, incorporando o valor prognóstico e preditivo da biologia tumoral (HORTOBAGYI *et al.*, 2017).

Históricamente, houve uma evolução nos critérios de decisão terapêutica, partindo do estadiamento clínico-anatômico e, com o avanço do entendimento sobre os subtipos biológicos, o foco expandiu-se para o estadiamento prognóstico clínico, integrando biomarcadores essenciais (RE, RP, HER2 e grau histológico). Mais recentemente, consolidou-se ainda o estadiamento prognóstico patológico, no qual as assinaturas genômicas podem ser utilizadas (HORTOBAGYI *et al.*, 2017; YOON *et al.*, 2019). Essa trajetória evidencia que o estadiamento do câncer de mama deixou de ser apenas uma medida anatômica, tornando-se um instrumento dinâmico ao integrar a biologia tumoral, refletindo de forma mais precisa o risco individual do paciente.

Um aspecto importante sobre os achados desta coorte foi a elevada proporção de óbitos sem recorrência, atuando como evento competidor, principalmente nos subtipos HER2 enriquecido e Luminal A, o que justifica a escolha metodológica pelo modelo de Fine-Gray e pela função de incidência cumulativa. Na abordagem tradicional, que utiliza o estimador de Kaplan-Meier para avaliar a probabilidade de recorrência, os óbitos por outras causas seriam tratados como censura não informativa¹⁰, o que poderia levar à superestimação da probabilidade real de recorrência (NOORDZIJ *et al.*, 2013; PUTTER; FIOCCO; GESKUS, 2007). Embora o modelo de Cox causa-específico, elaborado para a avaliação indireta da proporcionalidade dos riscos, tenha

¹⁰A censura não informativa, ou independente, pressupõe que o mecanismo que levou à censura do indivíduo é independente do mecanismo do evento de interesse (PRINJA; GUPTA; VERMA, 2010). Na análise de sobrevivência tradicional, assume-se que o indivíduo censurado continuaria a apresentar a mesma probabilidade de recorrência que os indivíduos que permaneceram em acompanhamento. Nessa perspectiva, esta premissa torna-se biologicamente implausível na ocorrência de óbitos por outras causas, uma vez que o óbito impede, de forma definitiva, a recorrência.

apresentado estimativas de efeito semelhantes às da modelagem principal de Fine-Gray (Tabela B2), a opção pela apresentação dos riscos médios de subdistribuição ($sHR_{\text{médio}}$) e das curvas de incidência cumulativa foi adequada para o propósito deste estudo, pois quantificou o impacto prognóstico dos subtipos moleculares em um cenário em que o óbito por outras causas é uma barreira concreta à ocorrência de recidiva.

Reconhece-se que o alto custo do perfil de assinaturas genômicas limita sua incorporação abrangente aos registros de câncer, principalmente no contexto brasileiro, entretanto, os achados apresentados nesta análise reforçam o posicionamento de outros autores de que os subtipos moleculares intrínsecos acrescentam valor prognóstico aos estudos sobre o câncer de mama (PARKER *et al.*, 2009; PRAT *et al.*, 2015).

Essa necessidade premente de incorporar a avaliação molecular à prática clínica obteve um avanço recente e expressivo no Brasil, com a publicação da Portaria do Ministério da Saúde nº 25, de 11 de maio de 2026 (BRASIL, 2026). A decisão de incorporar ao Sistema Único de Saúde (SUS) o teste de sequenciamento de nova geração (NGS) sinaliza um esforço governamental para alinhar-se à evolução do manejo e da estratificação do risco no câncer de mama (LOIBL *et al.*, 2024). O teste engloba a pesquisa de mutações nos genes BRCA 1 e 2 em células germinativas. Embora o teste incorporado pelo SUS não corresponda às assinaturas de expressão gênica utilizadas nesta pesquisa, sua incorporação representa um avanço importante na integração da informação molecular ao manejo clínico do câncer de mama. Ao democratizar o acesso à análise de DNA na linha de cuidado do câncer de mama, o Brasil começa a consolidar políticas públicas que reconhecem a biologia tumoral e a genética como componentes do prognóstico e do planejamento terapêutico.

5.2 A dinâmica temporal dos riscos em estudos oncológicos

Neste estudo, a hipótese de proporcionalidade dos riscos, pressuposto central dos modelos de risco de Cox e Fine-Gray, foi rejeitada para algumas variáveis incluídas nos modelos de recorrência (Apêndice B) e de sobrevivência (Apêndice C), como idade, sexo, grau histológico, subtipo molecular, estadiamento tumoral e ano do diagnóstico.

O comportamento dinâmico dos subtipos moleculares observado neste estudo é descrito por Rueda *et al.* (2019). De modo semelhante, Zhang *et al.* (2019) demonstraram que o efeito prognóstico de variáveis clínicas como idade ao diagnóstico, grau histológico e status linfonodal não se manteve constante ao longo do tempo.

É plausível que o efeito do ano do diagnóstico e do estadiamento tumoral não tenha sido constante ao longo do tempo nesta análise. Este achado pode ter refletido as mudanças nos critérios de classificação TNM/AJCC documentadas na metodologia deste estudo, com diferentes

edições ao longo do período avaliado. No estudo de Jin, Chen e Yu (2016), que compara as 7ª e 8ª edições do sistema [TNM/AJCC](#), os autores demonstram que pacientes classificados segundo critérios distintos apresentam diferenças significativas na estratificação prognóstica e na sobrevida, reforçando a ideia de que mudanças nos sistemas de estadiamento podem modificar o risco observado.

O efeito da variável sexo também não foi constante nesta análise (Tabela [C5](#)). Comportamento semelhante foi documentado por Wang *et al.* (2019), que, ao avaliarem a mortalidade após o diagnóstico de câncer de mama em homens e mulheres, identificaram efeitos tempo-dependentes em determinados estágios tumorais, com impacto predominantemente concentrado nos anos iniciais pós-diagnóstico. O achado desses autores sugere que o sexo pode atuar como modificador de risco em contextos clínicos específicos e que o seu impacto não se mantém constante ao longo do tempo.

Na análise oncológica, o efeito das variáveis apresenta comportamento dinâmico, podendo variar em intensidade e significância ao longo do tempo de seguimento. Bellera *et al.* (2010) demonstraram efeitos não constantes do grau histológico e do status hormonal em pacientes com câncer de mama. Eles afirmam ainda que ignorar os efeitos variáveis ao longo do tempo pode levar à perda de informações biológicas relevantes e mascarar padrões específicos da evolução da doença. Portanto, esses autores discutem amplamente caminhos metodológicos para avaliar se o pressuposto da proporcionalidade dos riscos foi violado. Em síntese, a literatura demonstra que a ausência de validação dos pressupostos e a adoção de estratégias alternativas para corrigir a violação podem levar a interpretações equivocadas dos riscos ([BELLERA *et al.*, 2010](#); [SCHEMPER; WAKOUNIG; HEINZE, 2009](#); [FINE; GRAY, 1999](#)).

5.3 Sobrevida global e específica no câncer de mama

A comparação dos fatores prognósticos (idade, sexo, estadiamento tumoral e ano do diagnóstico) entre o [RHC](#) Estadual e o [RCBP](#) Internacional evidenciou que a magnitude da associação entre esses fatores e a mortalidade em cinco anos variou entre as bases de dados. Como registros hospitalares e populacionais desempenham papéis distintos e complementares na vigilância do câncer ([PIÑEROS *et al.*, 2017](#); [PENA *et al.*, 2025](#)), esta análise buscou comparar o comportamento dos fatores prognósticos, e não a estrutura dos registros. Assim, as diferenças nas estimativas de risco observadas refletem as características das populações representadas por cada registro.

Ao ajustar os modelos por idade, sexo e ano do diagnóstico, o estadiamento tumoral foi o preditor mais fortemente associado à mortalidade em ambos os contextos (Tabelas [9](#), [10](#), [11](#) e [12](#)). O aumento progressivo dos riscos entre os estágios I e IV foi acompanhado por uma redução

acentuada das probabilidades ajustadas de sobrevida global e específica, evidenciando o gradiente prognóstico do estadiamento, compatível com a evolução clínica do câncer de mama. Esse comportamento é explicado pela própria biologia da doença: enquanto tumores diagnosticados em estágios iniciais permanecem restritos à mama ou apresentam comprometimento locorregional limitado, os tumores em estágio IV já apresentam disseminação metastática (DE VISSER; JOYCE, 2023; NOLAN; LINDEMAN; VISVADER, 2023), fator que reduz substancialmente o controle da doença. Consequentemente, a probabilidade de sobrevida diminui à medida que a doença progride.

O valor prognóstico do estadiamento nos desfechos do câncer de mama é demonstrado na literatura recente. Uma análise de dados populacionais norte-americanos entre 2010 e 2021 demonstrou que, apesar dos avanços na sobrevida de pacientes com câncer de mama em estágio IV, especialmente nos subtipos positivos para receptores hormonais, o estágio avançado da doença continuou a representar um risco elevado de óbito (AVILA *et al.*, 2026). Esses achados reforçam a importância do diagnóstico precoce como estratégia para melhorar o prognóstico do câncer de mama.

Maior probabilidade de sobrevida específica em relação à sobrevida global foi observada neste estudo. Como a sobrevida específica considera apenas os óbitos atribuídos ao câncer de mama, enquanto a sobrevida global incorpora toda a mortalidade, independentemente da causa, fatores como comorbidades e o envelhecimento podem contribuir para reduzir a sobrevida global. Essa interpretação é consistente com evidências de que a carga de comorbidades constitui um fator independente da mortalidade global em pacientes com câncer de mama (COOPER *et al.*, 2023).

No presente estudo, a idade foi um preditor independente de risco de morte, tanto no contexto populacional quanto no hospitalar. A relevância da idade como fator prognóstico no câncer de mama também tem sido demonstrada em outros estudos. Jackson *et al.* (2023), ao analisarem pacientes na Colúmbia Britânica, Canadá, por meio da base de dados da *Breast Cancer Outcomes Unit* (BCOU), identificaram pior sobrevida em pacientes muito jovens (< 35 anos) e em idosos (≥ 80 anos), mesmo após ajuste por subtipo, estágio e fatores de tratamento. Essas evidências destacam a importância da idade como componente da estratificação do risco de morte.

Em relação ao sexo, observou-se, no presente estudo, maior risco de morte por todas as causas ao longo de 5 anos no registro populacional. Desigualdades na mortalidade entre os sexos já foram documentadas em investigações anteriores (SCOMERSI *et al.*, 2021; VARZARU *et al.*, 2024). Diferentemente dos resultados obtidos nesta análise, o estudo conduzido por Scomersi *et al.* (2021) não encontrou diferenças significativas na sobrevida global entre homens e mulheres, já em relação à sobrevida específica, os homens apresentaram pior desfecho. A literatura aponta que o câncer de mama masculino é frequentemente diagnosticado em idades e em estágios mais avançados, em indivíduos com múltiplas comorbidades, o que pode contribuir para a maior

mortalidade por todas as causas observada neste grupo (ZHAO *et al.*, 2025). Esses fatores ajudam a explicar por que as diferenças entre os sexos tendem a ser mais evidentes na mortalidade global do que na mortalidade específica por câncer de mama.

Quanto ao efeito protetor associado ao ano do diagnóstico identificado nesta análise, essa variável deve ser interpretada como um marcador temporal das mudanças ocorridas ao longo do período observado. Essa interpretação é corroborada por evidências recentes na literatura internacional que demonstram melhorias progressivas na sobrevida ao longo das últimas décadas (KIM *et al.*, 2025; VERDUZCO-AGUIRRE *et al.*, 2025). Em âmbito nacional, ao avaliar uma coorte histórica de 125.005 pacientes no Estado de São Paulo, Verduzco-Aguirre *et al.* (2025) demonstraram que a sobrevida global mediana aumentou de 10,7 para 11,7 anos entre os períodos de 2000 a 2004 e de 2010 a 2014.

Como o efeito protetor do ano do diagnóstico se comportou de forma muito similar tanto no RHC Estadual quanto no RCBP Internacional, esse paralelismo indica que, apesar de contextos assistenciais distintos, ambos os registros capturaram melhora no prognóstico à medida que os anos avançaram. No entanto, esse efeito favorável não apaga as desigualdades que persistem no controle da doença. O trabalho de Kim *et al.* (2025) revela que o ritmo do progresso é assimétrico: enquanto países desenvolvidos atingem menores taxas de mortalidade global, em países com Índice de Desenvolvimento Humano (IDH) mais baixo entre os analisados, incluindo o Brasil, as taxas de mortalidade ainda são crescentes.

5.4 Mapeamento geográfico e desigualdades na distribuição de serviços oncológicos no Brasil

Os dados oriundos do RHC Nacional revelaram um perfil de acometimento pelo câncer de mama em idades mais jovens (Tabela 14). O trabalho de Silva, Thuler e Melo (2025), que utilizaram essa mesma fonte de dados, revelou que mulheres com idade <40 anos tiveram um aumento anual significativo na incidência de mais de 1,6%. Isso reforça a ideia de que há um perfil emergente de câncer de mama em idades mais jovens no Brasil. Resultados semelhantes foram relatados por Silva *et al.* (2025). Os autores observaram uma tendência de aumento no diagnóstico de mulheres jovens em todas as regiões do país no período de 2014 a 2022, com destaque para a região Norte.

O predomínio de tumores nos estágios II e III também foi uma característica do RHC Nacional. Este achado é consistente com os já descritos para o câncer de mama feminino (JOMAR *et al.*, 2023). Como nesta análise a amostra masculina foi proporcionalmente pequena (1,04%), torna-se incerto se o predomínio dos estágios II e III se aplica também aos homens. Diante disso, no câncer de mama masculino, os estudos relatam maior frequência de tumores em estágio

IV (THULER *et al.*, 2021). Esse cenário reforça o padrão documentado na literatura de que as populações dos países em desenvolvimento descobrem a doença quando os sintomas se tornam evidentes (BRAY *et al.*, 2024; XIONG *et al.*, 2025). Ademais, essa desigualdade entre os sexos reflete uma barreira cultural no comportamento dos homens quanto à busca por cuidados de saúde. Embora tenham avaliado outras topografias oncológicas, Silva, Oliveira e Curado (2024) evidenciaram essa vulnerabilidade ao constatarem que uma parcela importante dos diagnósticos de câncer em homens só é registrada no momento do óbito, o que sinaliza graves lacunas na oportunidade de captação precoce desse público no Brasil.

Os dados desta análise revelaram que as distâncias percorridas para o tratamento oncológico variam conforme a região e a complexidade da modalidade terapêutica (Tabela 16). A hormonioterapia apresentou a menor distância mediana de deslocamento, o que se justifica pela natureza do tratamento. Por ser uma terapia oral, o manejo clínico e a dispensação do medicamento exigem menor infraestrutura do que as demais terapias analisadas. Contudo, o fato de os pacientes ainda enfrentarem um deslocamento mediano de 74,6 km evidencia que a distribuição e o acompanhamento clínico permanecem concentrados em polos de alta complexidade, revelando barreiras à descentralização da assistência oncológica para a atenção primária, mesmo nas etapas de menor dependência de suporte tecnológico complexo (BRASIL, 2023; NOGUEIRA, 2003).

A imunoterapia, por sua vez, apresentou a maior barreira ao deslocamento. Esse achado refletiu a limitada disponibilidade e elevada centralização dessa modalidade nos estados de São Paulo e do Rio Grande do Sul no período avaliado (CÂNCER, 2014). Este achado também reflete o período em que as terapias-alvo, representadas pelo anticorpo monoclonal trastuzumabe, iniciaram sua incorporação no SUS (BRASIL, 2012). Mais recentemente, estudos sinalizaram mudanças expressivas no cenário da imunoterapia no Brasil, impulsionadas pela imunoterapia celular com CAR-T e pela intensificação da interação entre centros acadêmicos, hospitais e colaborações internacionais. Contudo, essa realidade ainda se mostra fortemente polarizada na região Sudeste (SILVA *et al.*, 2024). Uma análise sobre o acesso às terapias imuno-oncológicas no Brasil demonstrou que o alto custo e a limitada incorporação pelo SUS obrigam muitos pacientes a recorrer à judicialização para garantir tratamento (FERREIRA *et al.*, 2018). Isso demonstra que o acesso a imunoterapias de última geração no Brasil ainda se distancia do princípio da equidade (BRASIL, 2017).

Para as terapias convencionais (cirurgia, quimioterapia e radioterapia), as medianas de distância percorridas neste estudo variaram entre 81,8 e 88,2 km (Tabela 16), no entanto, as regiões Norte e Nordeste concentraram maiores trajetos, enquanto as regiões Sudeste e Sul apresentaram deslocamentos curtos (Figura 11). Esses resultados se alinham a estudos recentes (FONSECA *et al.*, 2022; SOUSA *et al.*, 2026). A concordância das distâncias de deslocamento aqui encontradas com o padrão de baixa distância delimitado por Sousa *et al.* (2026) demonstra a consolidação do atendimento periférico em zonas macro-urbanas. Por outro lado, a manutenção de trajetos mais longos nas regiões Norte e Nordeste, como documentado por Fonseca *et al.*

(2022), sugere que a equidade pretendida pelo [SUS](#) ainda esbarra na centralização tecnológica dos grandes polos de tratamento.

5.5 Registros de câncer e desafios estruturais

O principal diferencial deste trabalho foi a utilização de informações provenientes de registros oncológicos de diferentes naturezas (moleculares, hospitalares e populacionais), o que permitiu uma abordagem multidimensional dos desfechos do câncer de mama. Entretanto, a completude e padronização dos dados constituíram um dos principais desafios metodológicos deste estudo. Chen *et al.* (2023) ao avaliar dados populacionais e hospitalares, demonstraram que as diferenças estruturais na forma com que os registros são coletados e processados refletem nas estimativas de sobrevivência observadas. Essa fragilidade na integração de ecossistemas de dados nos serviços de saúde também é demonstrada no trabalho de Whitfield *et al.* (2025). Giusti *et al.* (2024) reforçam que a harmonização e o controle de qualidade rigoroso dos registros de câncer são fundamentais para que os dados possam, de fato, orientar efetivamente políticas públicas de saúde.

No contexto brasileiro, essa problemática ganha contornos mais complexos devido às desigualdades regionais e estruturais do sistema de saúde. Pena *et al.* (2025) destacam avanços na integração de técnicas de coleta passiva e automatizada a diferentes ferramentas tecnológicas em alguns registros de câncer brasileiros; contudo, isso não ocorre de maneira uniforme em todos os estados. Silva, Oliveira e Curado (2024) confirmam essas desigualdades na qualidade dos dados de câncer disponíveis no Brasil, destacando que atrasos na disponibilização desses dados podem comprometer as estimativas de incidência e de sobrevivência no câncer. A falta de preenchimento de variáveis no [RHC](#) Nacional também já foi documentada em estudos semelhantes ([GRASSI *et al.*, 2024](#)).

Existe uma necessidade crítica apontada na literatura por dados de alta qualidade e integração tecnológica dos registros de câncer, para que se possa mapear o real impacto da doença em diferentes populações ([SILVA; OLIVEIRA; CURADO, 2024](#); [PERES *et al.*, 2016](#); [PIÑEROS *et al.*, 2021](#)). Nesse sentido, Edwards *et al.* (2022) demonstraram que a capacidade dos registros de câncer de cumprir de forma consistente os padrões de qualidade e pontualidade dos dados está diretamente atrelada ao nível do reporte eletrônico e à conformidade de notificação das fontes hospitalares. Os autores destacam que os registros que atingem o patamar máximo de excelência em vigilância do câncer apresentam maiores taxas de envio eletrônico automatizado de dados, com destaque para a integração de laudos laboratoriais e relatórios hospitalares. Além disso, estes registros se beneficiam de uma infraestrutura robusta de pessoal especializado, incluindo profissionais de tecnologia da informação e registradores de tumores certificados para o controle de qualidade. Os achados desses autores evidenciam que a modernização tecnológica e a transição

para a coleta de dados em tempo real não são apenas melhorias processuais, mas também pilares para garantir a qualidade analítica das pesquisas sobre câncer de mama.

Em última análise, a disponibilização pública de grandes repositórios de dados populacionais, hospitalares e genômicos atua como elemento central para a colaboração científica e o avanço da pesquisa oncológica, permitindo que pesquisadores ao redor do mundo validem suas hipóteses sem os custos da coleta primária. Ao mitigar as barreiras institucionais e geográficas, os dados de acesso aberto e de qualidade potencializam a vigilância do câncer, acelerando substancialmente a produção de conhecimento na oncologia.

6 LIMITAÇÕES DO ESTUDO

A utilização de grandes bases de dados secundárias e retrospectivas impõe limitações estruturais e residuais que devem ser consideradas na interpretação dos resultados. O delineamento observacional da pesquisa, por exemplo, não permite estabelecer de forma inequívoca nexos causais diretos entre as variáveis explicativas e os desfechos oncológicos.

Devido à natureza das fontes públicas consultadas, não foi possível acessar ou controlar variáveis clínicas que influenciam o prognóstico, como a presença de comorbidades e detalhes sobre a adesão exata aos protocolos terapêuticos. Adicionalmente, o estudo está sujeito ao viés de confundimento residual (CHU *et al.*, 2020).

A ausência de variáveis temporais detalhadas, como o ano do diagnóstico na coorte molecular, impediu o controle do efeito do tempo. Os métodos de rastreamento e manejo terapêutico do câncer de mama evoluíram ao longo dos anos e, embora as modalidades terapêuticas não tenham sido o objeto direto de avaliação deste modelo, a evolução histórica do rastreamento e dos tratamentos atua como um fator de confusão não mensurado nessa coorte (VAN MAAREN *et al.*, 2019).

A estratificação das análises, combinando os subtipos moleculares intrínsecos e as janelas temporais de seguimento, resultou em subgrupos com escassez de eventos. Consequentemente, as estimativas obtidas nesses estratos devem ser interpretadas sob uma perspectiva estritamente exploratória, haja vista a instabilidade matemática e a elevada incerteza estatística decorrentes do exaurimento amostral.

A modelagem multivariada da sobrevida global e específica em 5 anos no RHC Estadual e no RCBP Internacional enfrentou uma limitação estrutural decorrente da distinção cronológica entre os sistemas de estadiamento tumoral utilizados nestas bases de dados. A consequência direta dessa limitação é a possibilidade de viés de classificação (JIN; CHEN; YU, 2016).

O RHC Nacional apresentou lacunas informacionais específicas que restringiram o escopo

analítico. A abordagem estritamente descritiva preservou a utilidade dos dados, mas impediu a validação de modelos analíticos de sobrevida em nível nacional. As análises geográficas e de fluxo assistencial implementadas no [RHC](#) Nacional apresentam as limitações intrínsecas à anonimização dos dados. Embora cada registro corresponda a um evento clínico real que demandou suporte assistencial, a impossibilidade de identificar indivíduos distintos impediu determinar o número absoluto de indivíduos que se deslocaram para tratamento. A indisponibilidade do endereço residencial exato impediu a análise do trajeto de indivíduos cujo município de domicílio coincidia com o de tratamento, o que subestima as barreiras de acessibilidade intraurbana.

A inclusão de registros do sexo masculino nas análises descritivas e multivariadas assegurou a representatividade dessa população na amostra. Todavia, a baixa prevalência do câncer de mama em homens limita a generalização das conclusões ao público masculino.

Por fim, destaca-se que a delimitação do período de observação entre 2000 e 2015, embora metodologicamente necessária para viabilizar um horizonte de seguimento robusto, resultou em um anacronismo terapêutico em relação à prática clínica contemporânea. A baixa utilização de imunoterapia registrada no [RHC](#) Nacional reflete esse recuo histórico. Consequentemente, os resultados deste estudo devem ser interpretados como um mapeamento de uma coorte histórica.

7 CONCLUSÃO

Este estudo cumpriu o propósito de explorar o potencial de diferentes bases de dados de acesso público para investigar características demográficas, tumorais e terapêuticas do câncer de mama. A utilização de registros hospitalares, populacionais e moleculares permitiu ampliar a compreensão da doença sob perspectivas complementares, evidenciando como diferentes fontes de informação respondem a questões distintas, mas convergentes sobre o câncer de mama.

Diante dos objetivos propostos e dos resultados discutidos conclui-se que:

- Objetivos 1, 2 e 3: a caracterização dos perfis das coortes demonstrou que diferentes bases de dados oferecem informações complementares sobre o câncer de mama. Entre elas, os subtipos intrínsecos caracterizados por [PAM50](#) destacam-se pelo seu valor prognóstico para recorrência, cujo comportamento mostrou-se dinâmico ao longo do seguimento.
- Objetivos 4 e 5: a análise dos fatores prognósticos associados ao risco de morte e probabilidades de sobrevida demonstrou que o estadiamento tumoral permaneceu como principal determinante prognóstico, independentemente da natureza do registro analisado. Embora a magnitude das associações tenha variado entre as bases de dados, o comportamento dos fatores prognósticos mostrou-se consistente nos diferentes contextos.
- Objetivo 6: marcantes desigualdades refletem a epidemiologia do câncer de mama no

cenário brasileiro. O fluxo de deslocamento e o mapeamento dos polos de tratamento evidenciaram desigualdades na distribuição que impuseram deslocamentos intermunicipais e interestaduais a mais da metade da amostra no período estudado.

Por fim, esse estudo demonstra que compreender o câncer de mama exige analisar simultaneamente diferentes dimensões da doença. Essa abordagem amplia as possibilidades de utilização dos registros públicos, oferecendo subsídios para pesquisadores e gestores de saúde monitorarem padrões da doença, identificarem desigualdades assistenciais e planejarem estratégias voltadas ao aprimoramento da atenção oncológica.

Referências

ADAMS, S. A. *et al.* An Investigation of Travel Distance and Timeliness of Breast Cancer Treatment Among a Diverse Cohort in the United States. **International Journal of Environmental Research and Public Health**, v. 22, n. 2, p. 176, jan. 2025. DOI: [10.3390/ijerph22020176](https://doi.org/10.3390/ijerph22020176).

ALLISON, K. H. Prognostic and predictive parameters in breast pathology: a pathologist's primer. **Modern Pathology**, v. 34, p. 94–106, 2021. DOI: [10.1038/s41379-020-00704-7](https://doi.org/10.1038/s41379-020-00704-7).

ANDRE, F. *et al.* Biomarkers for Adjuvant Endocrine and Chemotherapy in Early-Stage Breast Cancer: ASCO Guideline Update. **Journal of Clinical Oncology**, v. 40, n. 16, p. 1816–1837, 2022. DOI: [10.1200/JCO.22.00069](https://doi.org/10.1200/JCO.22.00069).

APHALE, R. *et al.* Why Should We Avoid Excisional Biopsy of Breast Lump? ASOMA Guide for Managing a Discrete Breast Lump. **Indian Journal of Surgical Oncology**, v. 14, n. 1, p. 194–198, mar. 2023. DOI: [10.1007/s13193-022-01687-4](https://doi.org/10.1007/s13193-022-01687-4).

ARUN, B. *et al.* BRCA-mutated breast cancer: the unmet need, challenges and therapeutic benefits of genetic testing. **British Journal of Cancer**, v. 131, n. 9, p. 1400–1414, nov. 2024. DOI: [10.1038/s41416-024-02827-z](https://doi.org/10.1038/s41416-024-02827-z).

AUSTIN, P. C.; ALLIGNOL, A.; FINE, J. P. The number of primary events per variable affects estimation of the subdistribution hazard competing risks model. **Journal of Clinical Epidemiology**, v. 83, p. 75–84, 2017. DOI: [10.1016/j.jclinepi.2016.11.017](https://doi.org/10.1016/j.jclinepi.2016.11.017).

AVILA, J. *et al.* Stage IV Breast Cancer Incidence and Survival, 2010-2021. **JAMA Network Open**, v. 9, n. 5, e2612042, maio 2026. DOI: [10.1001/jamanetworkopen.2026.12042](https://doi.org/10.1001/jamanetworkopen.2026.12042).

BARRIENTOS-TORO, E. N.; DING, Q.; RASO, M. G. Translational Aspects in Metaplastic Breast Carcinoma. **Cancers**, v. 16, n. 7, p. 1433, abr. 2024. DOI: [10.3390/cancers16071433](https://doi.org/10.3390/cancers16071433).

BELLERA, C. A. *et al.* Variables with time-varying effects and the Cox model: Some statistical concepts illustrated with a prognostic factor study in breast cancer. **BMC Medical Research Methodology**, v. 10, n. 1, p. 20, dez. 2010. DOI: [10.1186/1471-2288-10-20](https://doi.org/10.1186/1471-2288-10-20).

BHARDWAJ, P. V. *et al.* Male Breast Cancer: a Review on Diagnosis, Treatment, and Survivorship. **Current Oncology Reports**, v. 26, n. 1, p. 34–45, jan. 2024. DOI: [10.1007/s11912-023-01489-z](https://doi.org/10.1007/s11912-023-01489-z).

BISWAS, S. K. *et al.* The Mammary Gland: Basic Structure and Molecular Signaling during Development. **International Journal of Molecular Sciences**, v. 23, n. 7, p. 3883, mar. 2022. DOI: [10.3390/ijms23073883](https://doi.org/10.3390/ijms23073883).

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002. Institui o Código Civil.** [S. l.: s. n.], 2002. p. 1. Diário Oficial da União: Brasília – DF. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/110406.htm. Acesso em: 13 out. 2025.

BRASIL. **Lei nº 14.758, de 19 de dezembro de 2023. Institui a Política Nacional de Prevenção e Controle do Câncer no âmbito do Sistema Único de Saúde (SUS).** [S. l.: s. n.], 2023. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2023/lei/114758.htm. Acesso em: 2 jun. 2026.

BRASIL. **Portaria de Consolidação nº 1, de 28 de setembro de 2017. Consolida as normas sobre os direitos e deveres dos usuários da saúde, da organização e do funcionamento do Sistema Único de Saúde.** [S. l.: s. n.], 2017. Disponível em: https://portalsinan.saude.gov.br/images/documentos/Legislacoes/Portaria_Consolidacao_1_28_SETEMBRO_2017.pdf. Acesso em: 2 jun. 2026.

BRASIL. **Portaria IBGE nº 1.098, de 27 de agosto de 2025: Estimativas da população para estados e municípios em 1º de julho de 2025.** [S. l.: s. n.], 2025. Diário Oficial da União, Seção 1, nº 163, p. 97. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html>. Acesso em: 29 set. 2025.

BRASIL. **Portaria nº 140, de 27 de fevereiro de 2014.** [S. l.: s. n.], 2014. Disponível em: <https://www.inca.gov.br/publicacoes/legislacao/portaria-140-27-fev-2014>. Acesso em: 1 out. 2025.

BRASIL. **Portaria nº 18, de 25 de julho de 2012. Torna pública a decisão de incorporar o trastuzumabe para o tratamento do câncer de mama inicial e avançado no âmbito do Sistema Único de Saúde (SUS).** [S. l.: s. n.], 2012. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/sctie/2012/prt0018_25_07_2012.html. Acesso em: 2 jun. 2026.

BRASIL. **Portaria SCTIE/MS nº 25, de 11 de maio de 2026: Torna pública a decisão de incorporar, no âmbito do Sistema Único de Saúde - SUS, o sequenciamento de nova geração (NGS) para identificação de mutação nos genes BRCA1/2 em mulheres com câncer de mama.** [S. l.: s. n.], 2026. Diário Oficial da União, Seção 1, Edição 88, p. 137. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-sctie/ms-n-25-de-11-de-maio-de-2026-705078831>. Acesso em: 11 jun. 2026.

BRASIL. **Protocolo clínico e diretrizes terapêuticas: câncer de mama.** Brasília, DF: Ministério da Saúde, 2024. Disponível em: https://www.gov.br/conitec/pt-br/midias/consultas/relatorios/2024/RRPCDTCnecderMama_CP.pdf. Acesso em: 8 out. 2025.

BRAWLEY, O. W.; LANSEY, D. G. Disparities in Breast Cancer Outcomes and How to Resolve Them. **Hematology/Oncology Clinics of North America**, v. 37, n. 1, p. 1–15, fev. 2023. DOI: [10.1016/j.hoc.2022.08.002](https://doi.org/10.1016/j.hoc.2022.08.002).

BRAY, F. *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**, v. 74, n. 3, p. 229–263, 2024. DOI: [10.3322/caac.21834](https://doi.org/10.3322/caac.21834).

BRISKEN, C.; N; O'MALLEY, B. Hormone Action in the Mammary Gland. **Cold Spring Harbor Perspectives in Biology**, v. 2, n. 12, a003178, dez. 2010. DOI: [10.1101/cshperspect.a003178](https://doi.org/10.1101/cshperspect.a003178).

CÂNCER, REDE. Um novo tratamento contra o câncer de próstata? Vacina terapêutica brasileira apresenta resultados promissores, mas precisa de mais testes antes de ser comercializada. [S. l.: s. n.], 2014. Disponível em: https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document/rrc-26-versao-integral_0.pdf. Acesso em: 2 jun. 2026.

CARDOSO, F. *et al.* Characterization of male breast cancer: results of the EORTC 10085/TBCRC/BIG/NABCG International Male Breast Cancer Program. **Annals of Oncology**, v. 29, n. 2, p. 405–417, fev. 2018. DOI: [10.1093/annonc/mdx651](https://doi.org/10.1093/annonc/mdx651).

CAREY, L. A. *et al.* Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. **JAMA**, p. 2492–2502, jun. 2006. DOI: [10.1001/jama.295.21.2492](https://doi.org/10.1001/jama.295.21.2492).

CARVALHO, F. M. Triple-negative breast cancer: from none to multiple therapeutic targets in two decades. **Frontiers in Oncology**, v. 13, p. 1244781, nov. 2023. DOI: [10.3389/fonc.2023.1244781](https://doi.org/10.3389/fonc.2023.1244781).

CASWELL-JIN, J. L. *et al.* Analysis of Breast Cancer Mortality in the US—1975 to 2019. **JAMA**, v. 331, n. 3, p. 233–241, jan. 2024. DOI: [10.1001/jama.2023.25881](https://doi.org/10.1001/jama.2023.25881).

CENTERS FOR DISEASE CONTROL AND PREVENTION. **How Cancer Registries Work**. [S. l.: s. n.], 2024. Disponível em: <https://www.cdc.gov/national-program-cancer-registries/about/how-cancer-registries-work.html>. Acesso em: 1 out. 2025.

CHARLOT, M. *et al.* Pathologies of the male breast. **Diagnostic and Interventional Imaging**, v. 94, n. 1, p. 26–37, jan. 2013. DOI: [10.1016/j.diii.2012.10.011](https://doi.org/10.1016/j.diii.2012.10.011).

CHAS, M. *et al.* Clinicopathologic predictors of lymph node metastasis in breast cancer patients according to molecular subtype. **Journal of Gynecology Obstetrics and Human Reproduction**, p. 9–15, jan. 2018. DOI: <https://doi.org/10.1016/j.jogoh.2017.10.008>.

CHE, W. *et al.* How to use the Surveillance, Epidemiology, and End Results (SEER) data: research design and methodology. **Military Medical Research**, v. 10, n. 1, p. 50, 2023. DOI: [10.1186/s40779-023-00488-2](https://doi.org/10.1186/s40779-023-00488-2).

CHEANG, M. C. U. *et al.* Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. **JNCI: Journal of the National Cancer Institute**, v. 101, n. 10, p. 736–750, maio 2009. DOI: [10.1093/jnci/djp082](https://doi.org/10.1093/jnci/djp082).

CHEN, J. G. *et al.* Cancer survival: left truncation and comparison of results from hospital-based cancer registry and population-based cancer registry. **Frontiers in Oncology**, v. 13, p. 1173828, jun. 2023. DOI: [10.3389/fonc.2023.1173828](https://doi.org/10.3389/fonc.2023.1173828).

CHEN, L. *et al.* Locoregional recurrence in studies of primary systemic therapy in early invasive breast cancer. **The Breast**, v. 77, p. 103791, out. 2024. DOI: [10.1016/j.breast.2024.103791](https://doi.org/10.1016/j.breast.2024.103791).

CHIDAMBARAM, A. *et al.* Male Breast Cancer: Current Scenario and Future Perspectives. **Technology in Cancer Research & Treatment**, v. 23, p. 15330338241261836, jan. 2024. DOI: [10.1177/15330338241261836](https://doi.org/10.1177/15330338241261836).

CHO, S. B. Uncovering Oncogenic Mechanisms of Tumor Suppressor Genes in Breast Cancer Multi-Omics Data. **International Journal of Molecular Sciences**, v. 23, n. 17, p. 9624, ago. 2022. DOI: [10.3390/ijms23179624](https://doi.org/10.3390/ijms23179624).

CHU, L. *et al.* Consideration of confounding in epidemiologic studies assessing alcohol consumption on the risk of breast cancer: A brief report. **Chemico-Biological Interactions**, v. 322, p. 109060, maio 2020. DOI: [10.1016/j.cbi.2020.109060](https://doi.org/10.1016/j.cbi.2020.109060).

COATES, A. S. *et al.* Tailoring therapies—improving the management of early breast cancer: St Gallen international expert consensus on the primary therapy of early breast cancer 2015. **Annals of Oncology**, v. 26, n. 8, p. 1533–1546, 2015. DOI: [10.1093/annonc/mdv221](https://doi.org/10.1093/annonc/mdv221).

CONTRIBUTORS, OPENSTREETMAP. **OpenStreetMap Project**. [S. l.: s. n.], 2025. Disponível em: <https://www.openstreetmap.org>. Acesso em: 24 maio 2026.

COOPER, R. M. *et al.* Influence of Comorbidity Burden, Socioeconomic Status, and Race and Ethnicity on Survival Disparities in Patients With Cancer. **Cancer Control**, v. 30, p. 10732748231204474, 2023. DOI: [10.1177/10732748231204474](https://doi.org/10.1177/10732748231204474).

COSAR, R. *et al.* Breast Cancer Subtypes and Prognosis: Answers to Subgroup Classification Questions, Identifying the Worst Subgroup in Our Single-Center Series. **Breast Cancer: Targets and Therapy**, v. 14, p. 259–280, set. 2022. DOI: [10.2147/BCTT.S380754](https://doi.org/10.2147/BCTT.S380754).

CUCCINIELLO, L. *et al.* Tailoring adjuvant endocrine therapy in early breast cancer: When, how, and how long? **Cancer Treatment Reviews**, v. 110, p. 102445, nov. 2022. DOI: [10.1016/j.ctrv.2022.102445](https://doi.org/10.1016/j.ctrv.2022.102445).

CURTIS, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. **Nature**, v. 486, n. 7403, p. 346–352, 2012. DOI: [10.1038/nature10983](https://doi.org/10.1038/nature10983).

DA SILVA, J. L.; THULER, L. C. S.; DE MELO, A. C. Breast cancer patterns by age groups in Brazil: insights from population-based registries data. **BMC Cancer**, v. 25, n. 1, p. 18, 2025. DOI: [10.1186/s12885-024-13381-5](https://doi.org/10.1186/s12885-024-13381-5).

DALLA, Erica *et al.* Dormancy in Breast Cancer. **Cold Spring Harbor Perspectives in Medicine**, v. 13, n. 11, a041331, nov. 2023. DOI: [10.1101/cshperspect.a041331](https://doi.org/10.1101/cshperspect.a041331).

DE HAAS, S. L. *et al.* Tumor biomarkers and efficacy in patients treated with trastuzumab emtansine + pertuzumab versus standard of care in HER2-positive early breast cancer: an open-label, phase III study (KRISTINE). **Breast Cancer Research**, v. 25, n. 1, p. 2, jan. 2023. DOI: [10.1186/s13058-022-01587-z](https://doi.org/10.1186/s13058-022-01587-z).

DE VISSER, K. E.; JOYCE, J. A. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. **Cancer Cell**, v. 41, n. 3, p. 374–403, mar. 2023. DOI: [10.1016/j.ccell.2023.02.016](https://doi.org/10.1016/j.ccell.2023.02.016).

DUGGAN, M. A. *et al.* The Surveillance, Epidemiology, and End Results (SEER) Program and Pathology: Toward Strengthening the Critical Relationship. **American Journal of Surgical Pathology**, v. 40, n. 12, e94–e102, dez. 2016. DOI: [10.1097/PAS.0000000000000749](https://doi.org/10.1097/PAS.0000000000000749).

EDWARDS, P. *et al.* Características Operacionais dos Registros Centrais de Câncer que Apoiam a Geração de Dados de Vigilância de Alta Qualidade. **Gestão do Registro J.**, v. 49, n. 1, p. 10–16, 2022. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10036081/>. Acesso em: 2 jun. 2026.

FAN, T. *et al.* Therapeutic cancer vaccines: advancements, challenges and prospects. **Signal Transduction and Targeted Therapy**, v. 8, n. 1, p. 450, dez. 2023. DOI: [10.1038/s41392-023-01674-3](https://doi.org/10.1038/s41392-023-01674-3).

FERREIRA, C. G. *et al.* Increasing access to immuno-oncology therapies in Brazil. **Journal of Cancer Policy**, v. 16, p. 1–5, jun. 2018. DOI: [10.1016/j.jcipo.2017.12.004](https://doi.org/10.1016/j.jcipo.2017.12.004).

FINE, J. P.; GRAY, R. J. A proportional hazards model for the subdistribution of a competing risk. **Journal of the American Statistical Association**, v. 94, n. 446, p. 496–509, 1999. DOI: [10.1080/01621459.1999.10474144](https://doi.org/10.1080/01621459.1999.10474144).

FONSECA, B. P. *et al.* Geographic accessibility to cancer treatment in Brazil: A network analysis. **The Lancet Regional Health - Americas**, v. 7, p. 100153, mar. 2022. DOI: [10.1016/j.lana.2021.100153](https://doi.org/10.1016/j.lana.2021.100153).

FOUGNER, C. *et al.* Re-definition of claudin-low as a breast cancer phenotype. **Nature Communications**, v. 11, n. 1, p. 1787, abr. 2020. DOI: [10.1038/s41467-020-15574-5](https://doi.org/10.1038/s41467-020-15574-5).

FOX, J.; MONETTE, G. Generalized Collinearity Diagnostics. **Journal of the American Statistical Association**, v. 87, n. 417, p. 178–183, mar. 1992. DOI: [10.2307/2290467](https://doi.org/10.2307/2290467).

FRIEDMAN, S.; NEGOITA, S. History of the Surveillance, Epidemiology, and End Results (SEER) Program. **JNCI Monographs**, v. 2024, n. 65, p. 105–109, ago. 2024. DOI: [10.1093/jncimonographs/lgae033](https://doi.org/10.1093/jncimonographs/lgae033).

FUNDAÇÃO ONCOCENTRO DE SÃO PAULO. **Dicionário de Dados do RHC/SP**. [S. l.: s. n.], 2024. Disponível em: <https://fosp.saude.sp.gov.br/download/dicionario-de-dados/>. Acesso em: 14 maio 2026.

GAO, J. *et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. **Science Signaling**, v. 6, n. 269, abr. 2013. DOI: [10.1126/scisignal.2004088](https://doi.org/10.1126/scisignal.2004088).

GAO, J. J.; SWAIN, S. M. Luminal A Breast Cancer and Molecular Assays: A Review. **The Oncologist**, v. 23, n. 5, p. 556–565, maio 2018. DOI: [10.1634/theoncologist.2017-0535](https://doi.org/10.1634/theoncologist.2017-0535).

GENTILINI, O. D. *et al.* Sentinel Lymph Node Biopsy vs No Axillary Surgery in Patients With Small Breast Cancer and Negative Results on Ultrasonography of Axillary Lymph Nodes: The SOUND Randomized Clinical Trial. **JAMA Oncology**, v. 9, n. 11, p. 1557, nov. 2023. DOI: [10.1001/jamaoncol.2023.3759](https://doi.org/10.1001/jamaoncol.2023.3759).

GIRAUD, T. *et al.* **osrm: Interface entre R e o serviço de roteamento baseado no OpenStreetMap OSRM**. [S. l.], 2025. R package version 5.0.0. DOI: [10.32614/CRAN.package.osrm](https://doi.org/10.32614/CRAN.package.osrm). Disponível em: <https://CRAN.R-project.org/package=osrm>. Acesso em: 24 maio 2026.

GIUSTI, F. *et al.* Facing further challenges in cancer data quality and harmonisation. **Frontiers in Oncology**, v. 14, p. 1438805, jul. 2024. DOI: [10.3389/fonc.2024.1438805](https://doi.org/10.3389/fonc.2024.1438805).

GOLDHIRSCH, A. *et al.* Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. **Annals of Oncology**, v. 22, n. 8, p. 1736–1747, 2011. DOI: [10.1093/annonc/mdr304](https://doi.org/10.1093/annonc/mdr304).

GOVERNMENT OF CANADA. **Canadian Cancer Registry (CCR)**. [S. l.: s. n.], 2025. Disponível em: <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3207>. Acesso em: 2 out. 2025.

GRAMBSCH, P. M.; THERNEAU, T. M. Proportional hazards tests and diagnostics based on weighted residuals. **Biometrika**, v. 81, n. 3, p. 515–526, 1994. DOI: [10.1093/biomet/81.3.515](https://doi.org/10.1093/biomet/81.3.515).

GRASSI, J. *et al.* Temporal Trends in the Completeness of Epidemiological Variables in a Hospital-Based Cancer Registry of a Pediatric Oncology Center in Brazil. **International Journal of Environmental Research and Public Health**, v. 21, n. 2, p. 200, fev. 2024. DOI: [10.3390/ijerph21020200](https://doi.org/10.3390/ijerph21020200).

GUIMARÃES, J. M. N. *et al.* Income Segregation, Conditional Cash Transfers, and Breast Cancer Mortality Among Women in Brazil. **JAMA Network Open**, v. 7, n. 1, e2353100, jan. 2024. DOI: [10.1001/jamanetworkopen.2023.53100](https://doi.org/10.1001/jamanetworkopen.2023.53100).

GUPTA, A. *et al.* Impact of environmental temperature on the survival outcomes of breast cancer: A SEER-based study. **Breast Cancer Research and Treatment**, v. 207, n. 2, p. 383–392, set. 2024. DOI: [10.1007/s10549-024-07369-9](https://doi.org/10.1007/s10549-024-07369-9).

HOHMANN, L. *et al.* Genomic characterization of the HER2-enriched intrinsic molecular subtype in primary ER-positive HER2-negative breast cancer. **Nature Communications**, v. 16, n. 1, p. 2208, mar. 2025. DOI: [10.1038/s41467-025-57419-z](https://doi.org/10.1038/s41467-025-57419-z).

HORTOBAGYI, G. N. *et al.* Breast. In: AMERICAN JOINT COMMITTEE ON CANCER. **AJCC Cancer Staging Manual**. 8. ed. Chicago: AJCC, 2017. p. 589–628. Disponível em: <http://www.breastsurgeonsweb.com/wp-content/uploads/downloads/2020/10/AJCC-Breast-Cancer-Staging-System.pdf>. Acesso em: 2 out. 2025.

HOXHA, I. *et al.* Breast Cancer and Lifestyle Factors. **Hematology/Oncology Clinics of North America**, v. 38, n. 1, p. 137–170, fev. 2024. DOI: [10.1016/j.hoc.2023.07.005](https://doi.org/10.1016/j.hoc.2023.07.005).

INSTITUTO NACIONAL DE CÂNCER. **RHCNet – Integrador RHC**. [S. l.: s. n.], 2026. Disponível em: <https://irhc.inca.gov.br/RHCNet/visualizaTabNetExterno.action>. Acesso em: 14 maio 2026.

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. IARC. **National Cancer Institute (INCA) – Brazil**. [S. l.: s. n.], 2023. IARC. Disponível em: <https://gicr.iarc.who.int/training-center/brazil-iarc-gicr-collaborating-centre/>. Acesso em: 29 set. 2025.

JACKSON, E. *et al.* Does age affect outcome with breast cancer? **The Breast**, v. 70, p. 25–31, ago. 2023. DOI: [10.1016/j.breast.2023.06.001](https://doi.org/10.1016/j.breast.2023.06.001).

JANIC, A.; ABAD, E.; AMELIO, I. Decoding p53 tumor suppression: a crosstalk between genomic stability and epigenetic control? **Cell Death & Differentiation**, v. 32, n. 1, p. 1–8, jan. 2025. DOI: [10.1038/s41418-024-01259-9](https://doi.org/10.1038/s41418-024-01259-9).

JAROENSRI, R. *et al.* Deep learning models for histologic grading of breast cancer and association with disease prognosis. **npj Breast Cancer**, v. 8, n. 1, p. 113, out. 2022. DOI: [10.1038/s41523-022-00478-y](https://doi.org/10.1038/s41523-022-00478-y).

JAVED, A.; LTEIF, A. Development of the Human Breast. **Seminars in Plastic Surgery**, v. 27, n. 01, p. 005–012, maio 2013. DOI: [10.1055/s-0033-1343989](https://doi.org/10.1055/s-0033-1343989).

JIANG, R. *et al.* Association of education level with the risk of female breast cancer: a prospective cohort study. **BMC Women's Health**, v. 23, n. 1, p. 91, mar. 2023. DOI: [10.1186/s12905-023-02245-y](https://doi.org/10.1186/s12905-023-02245-y).

JIN, Y.; CHEN, M.; YU, X. Comparison of the 7th and proposed 8th editions of the AJCC/UICC TNM staging system for non-small cell lung cancer undergoing radical surgery. **Scientific Reports**, v. 6, n. 1, p. 33587, set. 2016. DOI: [10.1038/srep33587](https://doi.org/10.1038/srep33587).

JOMAR, R. T. *et al.* Fatores associados ao tempo para submissão ao primeiro tratamento do câncer de mama. **Ciência & Saúde Coletiva**, v. 28, n. 7, p. 2155–2164, jul. 2023. DOI: [10.1590/1413-81232023287.14982022](https://doi.org/10.1590/1413-81232023287.14982022).

KCIUK, M. *et al.* Recent Advances in Molecular Mechanisms of Cancer Immunotherapy. **Cancers**, v. 15, n. 10, p. 2721, maio 2023. DOI: [10.3390/cancers15102721](https://doi.org/10.3390/cancers15102721).

KERR, A. J. *et al.* Adjuvant and neoadjuvant breast cancer treatments: A systematic review of their effects on mortality. **Cancer Treatment Reviews**, v. 105, p. 102375, abr. 2022. DOI: [10.1016/j.ctrv.2022.102375](https://doi.org/10.1016/j.ctrv.2022.102375).

KIM, J.; MUNSTER, P. N. Estrogens and breast cancer. **Annals of Oncology**, v. 36, n. 2, p. 134–148, fev. 2025. DOI: [10.1016/j.annonc.2024.10.824](https://doi.org/10.1016/j.annonc.2024.10.824).

KIM, Joanne *et al.* Global patterns and trends in breast cancer incidence and mortality across 185 countries. **Nature Medicine**, v. 31, n. 4, p. 1154–1162, abr. 2025. DOI: [10.1038/s41591-025-03502-3](https://doi.org/10.1038/s41591-025-03502-3).

KOYA, A. I.; IBRAHIM, S. A. Carcinogenesis. *In*: STATPEARLS [Internet]. Treasure Island (FL): StatPearls Publishing, 2025. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK604463/>. Acesso em: 17 fev. 2025.

KUMAR, A. J. *et al.* Development and validation of a prediction model for 1-year mortality among older adults with Hodgkin Lymphoma who receive dose-intense chemotherapy. **Journal of Geriatric Oncology**, v. 12, n. 8, p. 1233–1239, nov. 2021. DOI: [10.1016/j.jgo.2021.07.003](https://doi.org/10.1016/j.jgo.2021.07.003).

LEE, S. W.; LEE, H. Engineered T Cell Receptor for Cancer Immunotherapy. **Biomolecules & Therapeutics**, v. 32, n. 4, p. 424–431, jul. 2024. DOI: [10.4062/biomolther.2023.197](https://doi.org/10.4062/biomolther.2023.197).

LI, F. *et al.* Prognostic predictive value of Ki-67 in stage I–II triple-negative breast cancer. **Future Science OA**, v. 10, n. 1, fso936, dez. 2024. DOI: [10.3389/fonc.2024.1388869](https://doi.org/10.3389/fonc.2024.1388869).

LI, J.; WU, J.; HAN, J. Analysis of Tumor Microenvironment Heterogeneity among Breast Cancer Subtypes to Identify Subtype-Specific Signatures. **Genes**, v. 14, n. 1, p. 44, dez. 2022. DOI: [10.3390/genes14010044](https://doi.org/10.3390/genes14010044).

LI, Y. *et al.* Global, regional, and national burden of male breast cancer in 204 countries and territories: a systematic analysis from the global burden of disease study, 1990–2021. **eClinicalMedicine**, v. 80, fev. 2025. DOI: [10.1016/j.eclinm.2024.103027](https://doi.org/10.1016/j.eclinm.2024.103027).

LI, Z. H. *et al.* Luminal B breast cancer: patterns of recurrence and clinical outcome. **Oncotarget**, v. 7, n. 40, p. 65024–65033, out. 2016. DOI: [10.18632/oncotarget.11344](https://doi.org/10.18632/oncotarget.11344).

LOIBL, S. *et al.* Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. **Annals of Oncology**, v. 35, n. 2, p. 159–182, fev. 2024. DOI: [10.1016/j.annonc.2023.11.016](https://doi.org/10.1016/j.annonc.2023.11.016).

LUTENCO, V. *et al.* Innovative Surgical Approaches That Improve Individual Outcomes in Advanced Breast Cancer. **International Journal of Women's Health**, v. 16, p. 555–560, mar. 2024. DOI: [10.2147/IJWH.S447837](https://doi.org/10.2147/IJWH.S447837).

MARCELINO, A. C. *et al.* Lower breast cancer survival among Black women in Brazil: a population-based retrospective study. **Public Health**, v. 217, p. 190–195, abr. 2023. DOI: [10.1016/j.puhe.2023.02.004](https://doi.org/10.1016/j.puhe.2023.02.004).

MARKARIAN, S.; HOLMES, D. R. Mammary Paget's Disease: An Update. **Cancers**, v. 14, n. 10, p. 2422, maio 2022. DOI: [10.3390/cancers14102422](https://doi.org/10.3390/cancers14102422).

MARTINS, L. F. L. *et al.* Perfil Epidemiológico da Incidência de Câncer no Brasil e Regiões: Estimativas para o Triênio 2026–2028. **Revista Brasileira de Cancerologia**, v. 72, n. 2, fev. 2026. DOI: [10.32635/2176-9745.RBC.2026v72n2.5587](https://doi.org/10.32635/2176-9745.RBC.2026v72n2.5587).

MCSHERRY, E. A. *et al.* Common Molecular Mechanisms of Mammary Gland Development and Breast Cancer: Molecular basis of invasion in breast cancer. **Cellular and Molecular Life Sciences**, v. 64, n. 24, p. 3201–3218, dez. 2007. DOI: [10.1007/s00018-007-7388-0](https://doi.org/10.1007/s00018-007-7388-0).

MINISTÉRIO DA SAÚDE. **Registros Hospitalares de Câncer: Rotinas e Procedimentos**. 1. ed. Rio de Janeiro, Brasil: Ministério da Saúde. Instituto Nacional de Câncer. Coordenação Nacional de Controle do Tabagismo, Prevenção e Vigilância de Câncer, 2000. ISBN 85-7318-051-X.

MOHANTY, S. S. *et al.* Impact of Cumulative Exposure to Circulating Ovarian Sex Hormones on Increasing the Risk of Hormone Receptor-Positive Breast Cancer. **Journal of Bio-X Research**, v. 7, p. 0005, jan. 2024. DOI: [10.34133/jbioxresearch.0005](https://doi.org/10.34133/jbioxresearch.0005).

MUKHERJEE, Abhik *et al.* Associations between genomic stratification of breast cancer and centrally reviewed tumour pathology in the METABRIC cohort. **npj Breast Cancer**, v. 4, n. 1, p. 5, mar. 2018. DOI: [10.1038/s41523-018-0056-8](https://doi.org/10.1038/s41523-018-0056-8).

NATIONAL INSTITUTES OF HEALTH. **National Cancer Institute. Accessing the Data – SEER Datasets**. [S. l.: s. n.], n.d. NIH. Disponível em: <https://seer.cancer.gov/data/access.html>. Acesso em: 2 out. 2025.

NATIONAL INSTITUTES OF HEALTH. **National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER). Cancer Registry**. [S. l.: s. n.], n.d. NIH. Disponível

em: https://seer.cancer.gov/registries/cancer_registry/cancer_registry.html. Acesso em: 2 out. 2025.

NATIONAL INSTITUTES OF HEALTH. **National Cancer Institute. Types of Radiation Therapy. SEER Training.** [S. l.: s. n.], 2023. NIH. Disponível em: <https://training.seer.cancer.gov/treatment/radiation/radiation-types.html>. Acesso em: 9 out. 2025.

NETO, L. A. A. Quantificar, padronizar e planejar: registros de câncer e atenção oncológica no Brasil (1960-1980). **Caminhos da História**, p. 73–96, jul. 2021. DOI: [10.38049/issn.2317-0875v26n2p.73-96](https://doi.org/10.38049/issn.2317-0875v26n2p.73-96).

NOGUEIRA, V. M. R. A descentralizacao da saude, atencao oncologica e estrategias de inclusao. **Revista Katalysis**, v. 6, n. 1, p. 104–112, 2003. Disponível em: <http://www.redalyc.org/articulo.oa?id=179618281012>. Acesso em: 2 jun. 2026.

NOLAN, E.; LINDEMAN; G. J.; VISVADER, J. E. Deciphering breast cancer: from biology to the clinic. **Cell**, v. 186, n. 8, p. 1708–1728, abr. 2023. DOI: [10.1016/j.cell.2023.01.040](https://doi.org/10.1016/j.cell.2023.01.040).

NOORDZIJ, M. *et al.* When do we need competing risks methods for survival analysis in nephrology? **Nephrology Dialysis Transplantation**, v. 28, n. 11, p. 2670–2677, 2013. DOI: [10.1093/ndt/gft355](https://doi.org/10.1093/ndt/gft355).

O'BRIEN, R. M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. **Quality & Quantity**, v. 41, n. 5, p. 673–690, 2007. DOI: [10.1007/s11135-006-9018-6](https://doi.org/10.1007/s11135-006-9018-6).

OBEAGU, E. I.; OBEAGU, G. U. Breast cancer: A review of risk factors and diagnosis. **Medicine**, v. 103, n. 3, e36905, jan. 2024. DOI: [10.1097/MD.00000000000036905](https://doi.org/10.1097/MD.00000000000036905).

ORGANIZATION, WORLD HEALTH. **International Classification of Diseases for Oncology (ICD-O): First Revision.** 3. ed. Geneva: World Health Organization, 2013. Disponível em: <https://www.who.int/publications/b/31314>. Acesso em: 15 out. 2025.

ORRANTIA-BORUNDA, E. *et al.* Subtypes of Breast Cancer. In: MAYROVITZ, HARVEY N. (ed.). **Breast Cancer.** [S. l.]: Exon Publications, 2022. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK583808/>. Acesso em: 5 out. 2025.

PARK, K. U. *et al.* Sentinel Lymph Node Biopsy in Early-Stage Breast Cancer: ASCO Guideline Update. **Journal of Clinical Oncology**, v. 43, n. 14, p. 1720–1741, maio 2025. DOI: [10.1200/JCO-25-00099](https://doi.org/10.1200/JCO-25-00099).

PARKER, J. S. *et al.* Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. **Journal of Clinical Oncology**, v. 27, n. 8, p. 1160–1167, mar. 2009. DOI: [10.1200/JCO.2008.18.1370](https://doi.org/10.1200/JCO.2008.18.1370).

- PENA, G. P. *et al.* Registros de Câncer de Base Populacional no Brasil: Relevância, Desafios e Oportunidades. **Revista Brasileira de Cancerologia**, v. 71, n. 1, fev. 2025. DOI: [10.32635/2176-9745.RBC.2025v71n1.4878](https://doi.org/10.32635/2176-9745.RBC.2025v71n1.4878).
- PEREIRA, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. **Nature Communications**, v. 7, n. 1, p. 11479, 2016. DOI: [10.1038/ncomms11479](https://doi.org/10.1038/ncomms11479).
- PEREIRA, R. H. M. *et al.* geobr: Download Official Spatial Data Sets of Brazil. **CRAN Repository**, Version 2.0.0, 2026. DOI: [10.32614/CRAN.package.geobr](https://doi.org/10.32614/CRAN.package.geobr).
- PERES, S. V. *et al.* Melhora na qualidade e completitude da base de dados do Registro de Câncer de Base Populacional do município de São Paulo: uso das técnicas de linkage. **Revista Brasileira de Epidemiologia**, v. 19, n. 4, p. 753–765, dez 2016. DOI: [10.1590/1980-5497201600040006](https://doi.org/10.1590/1980-5497201600040006).
- PEROU, Charles M. *et al.* Molecular portraits of human breast tumours. **Nature**, v. 406, n. 6797, p. 747–752, ago. 2000. DOI: [10.1038/35021093](https://doi.org/10.1038/35021093).
- PIÑEROS, M. *et al.* A Global Cancer Surveillance Framework Within Noncommunicable Disease Surveillance: Making the Case for Population-Based Cancer Registries. **Epidemiologic Reviews**, v. 39, n. 1, p. 161–169, jan. 2017. DOI: [10.1093/epirev/mxx003](https://doi.org/10.1093/epirev/mxx003).
- PIÑEROS, M. *et al.* Progress, challenges and ways forward supporting cancer surveillance in Latin America. **International Journal of Cancer**, v. 149, n. 1, p. 12–20, jul. 2021. DOI: [10.1002/ijc.33407](https://doi.org/10.1002/ijc.33407).
- PLANEY, A. M. *et al.* The intersection of travel burdens and financial hardship in cancer care: a scoping review. **JNCI Cancer Spectrum**, v. 8, n. 5, pkae093, set. 2024. DOI: [10.1093/jncics/pkae093](https://doi.org/10.1093/jncics/pkae093).
- POMMIER, R. M. *et al.* Comprehensive characterization of claudin-low breast tumors reflects the impact of the cell-of-origin on cancer evolution. **Nature Communications**, v. 11, n. 1, p. 3431, jul. 2020. DOI: [10.1038/s41467-020-17249-7](https://doi.org/10.1038/s41467-020-17249-7).
- PRAT, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. **Breast Cancer Research**, v. 12, n. 5, r68, 2010. DOI: [10.1186/bcr2635](https://doi.org/10.1186/bcr2635).
- PRAT, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. **The Breast**, v. 24, s26–s35, nov. 2015. DOI: [10.1016/j.breast.2015.07.008](https://doi.org/10.1016/j.breast.2015.07.008).
- PRINJA, S.; GUPTA, N.; VERMA, R. Censoring in clinical trials: Review of survival analysis techniques. **Indian Journal of Community Medicine**, v. 35, n. 2, p. 217–221, abr. 2010. DOI: [10.4103/0970-0218.66859](https://doi.org/10.4103/0970-0218.66859).

PUTTER, H.; FIOCCO, M.; GESKUS, R. B. Tutorial in biostatistics: competing risks and multi-state models. **Statistics in Medicine**, v. 26, n. 11, p. 2389–2430, 2007. DOI: [10.1002/sim.2712](https://doi.org/10.1002/sim.2712).

QIU, Y. *et al.* Clinical Application of Cytokines in Cancer Immunotherapy. **Drug Design, Development and Therapy**, v. 15, p. 2269–2287, maio 2021. DOI: [10.2147/DDDT.S308578](https://doi.org/10.2147/DDDT.S308578).

QIU, Y. *et al.* Triple-negative breast cancer survival prediction: population-based research using the SEER database and an external validation cohort. **Frontiers in Oncology**, v. 14, p. 1388869, jun. 2024. DOI: [10.3389/fonc.2024.1388869](https://doi.org/10.3389/fonc.2024.1388869).

RAKHA, E.; TOSS, M.; QUINN, C. Specific cell differentiation in breast cancer: a basis for histological classification. **Journal of Clinical Pathology**, v. 75, n. 2, p. 76–84, fev. 2022. DOI: [10.1136/jclinpath-2021-207487](https://doi.org/10.1136/jclinpath-2021-207487).

RAMI-PORTA, R. The TNM classification of lung cancer—a historic perspective. **Journal of Thoracic Disease**, v. 16, n. 11, p. 8053–8067, nov. 2024. DOI: [10.21037/jtd-24-1212](https://doi.org/10.21037/jtd-24-1212).

RATHS, F. *et al.* The molecular consequences of androgen activity in the human breast. **Cell Genomics**, v. 3, n. 3, p. 100272, mar. 2023. DOI: [10.1016/j.xgen.2023.100272](https://doi.org/10.1016/j.xgen.2023.100272).

ROJAS-SALAZAR, Y. *et al.* Potential Drug Synergy Through the ERBB2 Pathway in HER2+ Breast Tumors. **International Journal of Molecular Sciences**, v. 25, n. 23, p. 12840, nov. 2024. DOI: [10.3390/ijms252312840](https://doi.org/10.3390/ijms252312840).

ROY, A. M. *et al.* A review of treatment options in HER2-low breast cancer and proposed treatment sequencing algorithm. **Cancer**, v. 129, n. 18, p. 2773–2788, set. 2023. DOI: [10.1002/cncr.34904](https://doi.org/10.1002/cncr.34904).

RUEDA, O. M. *et al.* Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. **Nature**, v. 567, n. 7748, p. 399–404, 2019. DOI: [10.1038/s41586-019-1007-8](https://doi.org/10.1038/s41586-019-1007-8).

SALEH, R. R. *et al.* Influence of Competing Risks on Estimates of Recurrence Risk and Breast Cancer-Specific Mortality in Analyses of the Early Breast Cancer Trialists' Collaborative Group. **Scientific Reports**, v. 10, n. 1, p. 4091, mar. 2020. DOI: [10.1038/s41598-020-61093-0](https://doi.org/10.1038/s41598-020-61093-0).

SCHEMPER, M.; WAKOUNIG, S.; HEINZE, G. The estimation of average hazard ratios by weighted Cox regression. **Statistics in Medicine**, v. 28, n. 19, p. 2473–2489, 2009. DOI: [10.1002/sim.3623](https://doi.org/10.1002/sim.3623).

SCOMERSI, S. *et al.* Comparison between male and female breast cancer survival using propensity score matching analysis. **Scientific Reports**, v. 11, n. 1, p. 11639, jun. 2021. DOI: [10.1038/s41598-021-91131-4](https://doi.org/10.1038/s41598-021-91131-4).

SEER PROGRAM. **SEER Research plus description. Cases diagnosed 1975-2021. SEER November 2023 Submission.** [S. l.]: National Cancer Institute, 2023. Disponível em: <https://seer.cancer.gov/data-software/documentation/seerstat/nov2023/TextData.FileDescription-nov2023.pdf>. Acesso em: 14 maio 2026.

SEVER, R.; BRUGGE, J. S. Signal Transduction in Cancer. **Cold Spring Harbor Perspectives in Medicine**, v. 5, n. 4, a006098–a006098, abr. 2015. DOI: [10.1101/cshperspect.a006098](https://doi.org/10.1101/cshperspect.a006098).

SHI, Z.; XIN, M. Endocrine Hormones and Their Impact on Pubertal Gynecomastia. **Journal of Clinical Medicine**, v. 14, n. 1, p. 158, dez. 2024. DOI: [10.3390/jcm14010158](https://doi.org/10.3390/jcm14010158).

SIEGEL, R. L. *et al.* Cancer Statistics, 2026. **CA: A Cancer Journal for Clinicians**, v. 76, n. 1, e70043, jan. 2026. DOI: [10.3322/caac.70043](https://doi.org/10.3322/caac.70043).

SILVA, C. C. *et al.* Emerging Activity of Cellular Immunotherapy for Treatment of Cancer in Brazil: Report from the Brazilian Registry. **Journal of Bone Marrow Transplantation and Cellular Therapy**, v. 5, n. 2, p. 235, jul. 2024. DOI: [10.46765/2675-374X.2024v5n2p235](https://doi.org/10.46765/2675-374X.2024v5n2p235).

SILVA, D. R. M.; OLIVEIRA, M. M.; CURADO, M. P. Data quality in Brazilian population-based cancer registries for gastrointestinal cancers. **BMC Cancer**, v. 24, n. 1, p. 870, jul. 2024. DOI: [10.1186/s12885-024-12477-2](https://doi.org/10.1186/s12885-024-12477-2).

SILVA, G. *et al.* Temporal Trend of Breast Cancer Diagnoses in Young Women in Different Regions of Brazil from 2014 to 2022. **Medical Research Archives**, v. 13, n. 3, 2025. DOI: [10.18103/mra.v13i3.6441](https://doi.org/10.18103/mra.v13i3.6441).

SLADE, E. *et al.* Health Inequalities in Breast Cancer in England: A Pragmatic Review to Inform National Institute for Care and Excellence (NICE) Recommendations. **Cancer Control**, v. 31, p. 10732748241293976, jan. 2024. DOI: [10.1177/10732748241293976](https://doi.org/10.1177/10732748241293976).

SONG, B.; SINGH, H. Rare Breast Cancers Review. **Healthcare**, v. 12, n. 23, p. 2483, dez. 2024. DOI: [10.3390/healthcare12232483](https://doi.org/10.3390/healthcare12232483).

SØRENSEN, H. T.; HORVÁTH-PUHÓ, E.; PEACOCK, J. L. An Introduction to Competing Risks in Epidemiology. **Clinical Epidemiology**, v. 18, p. 574928, fev. 2026. DOI: [10.2147/CLEP.S574928](https://doi.org/10.2147/CLEP.S574928).

SOUSA, E. L. *et al.* Geographic disparities and temporal trends regarding access to cancer treatment: a spatial analysis, Brazil, 2015-2022. **Epidemiologia e Serviços de Saúde**, v. 35, e20240725.a, 2026. DOI: [10.1590/S2237-96222026v35e20240725.a](https://doi.org/10.1590/S2237-96222026v35e20240725.a).

STALPERS, L. J. A.; KAPLAN, E. L. Edward L. Kaplan and the Kaplan-Meier Survival Curve. **BSHM Bulletin: Journal of the British Society for the History of Mathematics**, v. 33, n. 2, p. 109–135, maio 2018. DOI: [10.1080/17498430.2018.1450055](https://doi.org/10.1080/17498430.2018.1450055).

STANOWICKA-GRADA, M.; SENKUS, E. Anti-HER2 Drugs for the Treatment of Advanced HER2 Positive Breast Cancer. **Current Treatment Options in Oncology**, v. 24, n. 11, p. 1633–1650, nov. 2023. DOI: [10.1007/s11864-023-01137-5](https://doi.org/10.1007/s11864-023-01137-5).

TAKESHIMA, H.; USHIJIMA, T. Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. **npj Precision Oncology**, v. 3, n. 1, p. 1–8, mar. 2019. DOI: [10.1038/s41698-019-0079-0](https://doi.org/10.1038/s41698-019-0079-0).

TEXTOR, J. *et al.* Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. **International Journal of Epidemiology**, dyw341, jan. 2017. DOI: [10.1093/ije/dyw341](https://doi.org/10.1093/ije/dyw341).

THULER, L. C. S. *et al.* Characteristics and prognosis of male breast cancer in Brazil: A cohort study. **The Breast Journal**, v. 27, n. 1, p. 95–98, jan. 2021. DOI: [10.1111/tbj.14120](https://doi.org/10.1111/tbj.14120).

TINTERRI, C. *et al.* Sentinel lymph node biopsy versus axillary lymph node dissection in breast cancer patients undergoing mastectomy with one to two metastatic sentinel lymph nodes: sub-analysis of the SINODAR-ONE multicentre randomized clinical trial and reopening of enrolment. **British Journal of Surgery**, v. 110, n. 9, p. 1143–1152, ago. 2023. DOI: [10.1093/bjs/znad215](https://doi.org/10.1093/bjs/znad215).

VAN DER VEN, C. C. E. M. *et al.* Calculating follow-up completeness: a comparison of multiple methods under different simulated scenarios and a use case. **Journal of Clinical Epidemiology**, v. 182, p. 111757, 2025. DOI: [10.1016/j.jclinepi.2025.111757](https://doi.org/10.1016/j.jclinepi.2025.111757).

VAN MAAREN, M. C. *et al.* Different statistical techniques dealing with confounding in observational research: Measuring the effect of breast-conserving therapy and mastectomy on survival. **Journal of Cancer Research and Clinical Oncology**, v. 145, n. 6, p. 1485–1493, jun. 2019. DOI: [10.1007/s00432-019-02919-x](https://doi.org/10.1007/s00432-019-02919-x).

VARZARU, V. B. *et al.* Observational Study of Men and Women with Breast Cancer in Terms of Overall Survival. **Cancers**, v. 16, n. 17, p. 3049, set. 2024. DOI: [10.3390/cancers16173049](https://doi.org/10.3390/cancers16173049).

VERDUZCO-AGUIRRE, H. *et al.* Twenty years of breast cancer epidemiology and treatment patterns in São Paulo, Brazil—observed versus expected treatment utilization in a retrospective cohort. **The Lancet Regional Health - Americas**, v. 47, p. 101115, jul. 2025. DOI: [10.1016/j.lana.2025.101115](https://doi.org/10.1016/j.lana.2025.101115).

VERONA, F. *et al.* Targeting epigenetic alterations in cancer stem cells. **Frontiers in Molecular Medicine**, v. 2, p. 1011882, set. 2022. DOI: [10.3389/fmmed.2022.1011882](https://doi.org/10.3389/fmmed.2022.1011882).

VON ELM, E. *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. **The Lancet**, v. 370, n. 9596, p. 1453–1457, out. 2007. DOI: [10.1016/s0140-6736\(07\)61602-x](https://doi.org/10.1016/s0140-6736(07)61602-x).

VOUTSADAKIS, I. A. Comparison of Clinical Subtypes of Breast Cancer within the Claudin-Low Molecular Cluster Reveals Distinct Phenotypes. **Cancers**, v. 15, n. 10, p. 2689, maio 2023. DOI: [10.3390/cancers15102689](https://doi.org/10.3390/cancers15102689).

WANG, F.; others. Overall Mortality After Diagnosis of Breast Cancer in Men vs Women. **JAMA Oncology**, v. 5, n. 11, p. 1589–1594, nov. 2019. DOI: [10.1001/jamaoncol.2019.2803](https://doi.org/10.1001/jamaoncol.2019.2803).

WANG, J.; WU, S. Breast Cancer: An Overview of Current Therapeutic Strategies, Challenge, and Perspectives. **Breast Cancer: Targets and Therapy**, v. 15, p. 721–730, out. 2023. DOI: [10.2147/BCTT.S432526](https://doi.org/10.2147/BCTT.S432526).

WANG, X.; KATTAN, M. W. Cohort Studies. **Chest**, v. 158, n. 1, s72–s78, jul. 2020. DOI: [10.1016/j.chest.2020.03.014](https://doi.org/10.1016/j.chest.2020.03.014).

WHELAN, T. J. *et al.* Omitting Radiotherapy after Breast-Conserving Surgery in Luminal A Breast Cancer. **New England Journal of Medicine**, v. 389, n. 7, p. 612–619, ago. 2023. DOI: [10.1056/NEJMoa2302344](https://doi.org/10.1056/NEJMoa2302344).

WHITFIELD, E. *et al.* Differences in recording of cancer diagnosis between datasets in England: A population-based study of linked cancer registration, hospital, and primary care data. **Cancer Epidemiology**, v. 94, p. 102703, fev. 2025. DOI: [10.1016/j.canep.2024.102703](https://doi.org/10.1016/j.canep.2024.102703).

WU, Z.; XIA, F.; LIN, R. Global burden of cancer and associated risk factors in 204 countries and territories, 1980–2021: a systematic analysis for the GBD 2021. **Journal of Hematology & Oncology**, v. 17, n. 1, p. 119, nov. 2024. DOI: [10.1186/s13045-024-01640-8](https://doi.org/10.1186/s13045-024-01640-8).

XIONG, X. *et al.* Breast cancer: pathogenesis and treatments. **Signal Transduction and Targeted Therapy**, v. 10, n. 1, p. 49, fev. 2025. DOI: [10.1038/s41392-024-02108-4](https://doi.org/10.1038/s41392-024-02108-4).

YEKEDÜZ, E. *et al.* Comparison of Clinical and Pathological Factors Affecting Early and Late Recurrences in Patients with Operable Breast Cancer. **Journal of Clinical Medicine**, v. 11, n. 9, p. 2332, abr. 2022. DOI: [10.3390/jcm11092332](https://doi.org/10.3390/jcm11092332).

YIALLOUROU, A. *et al.* Non-genetic factors and breast cancer: an umbrella review of meta-analyses. **BMC Cancer**, v. 24, n. 1, p. 903, jul. 2024. DOI: [10.1186/s12885-024-12641-8](https://doi.org/10.1186/s12885-024-12641-8).

YOON, E. C. *et al.* Impact of biomarkers and genetic profiling on breast cancer prognostication: A comparative analysis of the 8th edition of breast cancer staging system. **The Breast Journal**, v. 25, p. 829–837, set. 2019. DOI: [10.1111/tbj.13352](https://doi.org/10.1111/tbj.13352).

ZHANG, M. *et al.* Time-varying effects of prognostic factors associated with long-term survival in breast cancer. **Endocrine-Related Cancer**, v. 25, n. 5, p. 509–521, maio 2018. DOI: [10.1530/ERC-17-0502](https://doi.org/10.1530/ERC-17-0502).

ZHAO, L. *et al.* Decoding male breast cancer: Epidemiological insights, cutting-edge treatments, and future perspectives. **Discover Oncology**, v. 16, n. 1, p. 360, 2025. DOI: [10.1007/s12672-025-02140-y](https://doi.org/10.1007/s12672-025-02140-y).

ZHU, K. *et al.* HER2-targeted therapies in cancer: a systematic review. **Biomarker Research**, v. 12, n. 1, p. 16, fev. 2024. DOI: [10.1186/s40364-024-00565-1](https://doi.org/10.1186/s40364-024-00565-1).

ZNAOR, A. *et al.* The International Classification of Diseases for Oncology, 4th Edition (ICD-O-4): An overview. **Cancer Epidemiology**, v. 101, p. 102989, abr. 2026. DOI: [10.1016/j.canep.2026.102989](https://doi.org/10.1016/j.canep.2026.102989).

ZOMBORI, T. *et al.* The prognostic value of histological grade determined after neoadjuvant chemotherapy of breast cancer. **Pathology - Research and Practice**, v. 265, p. 155732, 2025. DOI: [10.1016/j.prp.2024.155732](https://doi.org/10.1016/j.prp.2024.155732).

APÊNDICES E ANEXOS

Anexo A - Parecer ético

UNIVERSIDADE FEDERAL DE
UBERLÂNDIA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: ANÁLISE DA EFICÁCIA DE TRATAMENTOS PARA CÂNCER DE MAMA E FATORES DE RISCO ASSOCIADOS AO PROGNÓSTICO

Pesquisador: Marcelo José Barbosa Silva

Área Temática:

Versão: 1

CAAE: 86998025.3.0000.5152

Instituição Proponente: Instituto de Ciências Biomédicas

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 7.534.299

Apresentação do Projeto:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Objetivo da Pesquisa:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Avaliação dos Riscos e Benefícios:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Comentários e Considerações sobre a Pesquisa:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Considerações sobre os Termos de apresentação obrigatória:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Recomendações:

Vide campo "Conclusões ou Pendências e Lista de Inadequações".

Conclusões ou Pendências e Lista de Inadequações:

O CEP-UFU informa que não serão registradas nem avaliadas pelo sistema CEP/CONEP as pesquisas que utilizem informações de acesso público, nos termos da legislação vigente, como aquelas disponíveis em bases de dados públicas, desde que não identifiquem diretamente ou indiretamente os participantes.

Endereço: Av. João Naves de Ávila 2121 - Bloco "1A", sala 224 - Campus Sta. Mônica

Bairro: Santa Mônica

CEP: 38.408-144

UF: MG

Município: UBERLÂNDIA

Telefone: (34)3239-4131

Fax: (34)3239-4131

E-mail: cep@propp.ufu.br

Anexo A - Parecer ético

UNIVERSIDADE FEDERAL DE
UBERLÂNDIA



Continuação do Parecer: 7.534.299

O CEP/UFU destina-se a fazer a revisão ética de toda e qualquer pesquisa envolvendo o ser humano, de todas as áreas do conhecimento, sob a responsabilidade de um pesquisador e da instituição proponente à qual está vinculado. Segundo os termos e definições da Resolução nº 466/2012, "II.14 - pesquisa envolvendo seres humanos - pesquisa que, individual ou coletivamente, tenha como participante o ser humano, em sua totalidade ou partes dele, e o envolva de forma direta ou indireta, incluindo o manejo de seus dados, informações ou materiais biológicos." Este protocolo de pesquisa não envolve seres humanos.

Diante disso, o CEP-UFU indica o parecer de Retirado.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_2511331.pdf	13/03/2025 18:51:55		Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_CEP.pdf	13/03/2025 18:51:09	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	TERMO_DE_UTILIZACAO_DE_DADOS.pdf	13/03/2025 18:44:34	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	TERMO_DE_COMPROMISSO_EQUIPE_EXECUTORA.pdf	13/03/2025 18:17:08	GISELE CRISTIANE DA SILVA DIAS	Aceito
Folha de Rosto	FOLHA_DE_ROSTO.pdf	13/03/2025 14:37:01	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	JUSTIFICATIVA_AUSENCIA_TCLE.pdf	05/03/2025 19:41:22	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	INSTRUMENTO_DE_COLETA_DE_DADOS.pdf	05/03/2025 13:24:38	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	LATTES.pdf	05/03/2025 13:23:11	GISELE CRISTIANE DA SILVA DIAS	Aceito
Outros	LATTES_MARCELO.pdf	05/03/2025 13:21:56	GISELE CRISTIANE DA SILVA DIAS	Aceito

Situação do Parecer:

Retirado

Necessita Apreciação da CONEP:

Endereço: Av. João Naves de Ávila 2121 - Bloco "1A", sala 224 - Campus Sta. Mônica
Bairro: Santa Mônica **CEP:** 38.408-144
UF: MG **Município:** UBERLÂNDIA
Telefone: (34)3239-4131 **Fax:** (34)3239-4131 **E-mail:** cep@propp.ufu.br

Anexo A - Parecer ético

UNIVERSIDADE FEDERAL DE
UBERLÂNDIA



Continuação do Parecer: 7.534.299

Não

UBERLÂNDIA, 28 de Abril de 2025

Assinado por:
Eduardo Henrique Rosa Santos
(Coordenador(a))

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica
Bairro: Santa Mônica **CEP:** 38.408-144
UF: MG **Município:** UBERLÂNDIA
Telefone: (34)3239-4131 **Fax:** (34)3239-4131 **E-mail:** cep@propp.ufu.br

Anexo B - STROBE checklist

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses

Continued on next page

Anexo B - STROBE checklist

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —Report numbers of outcome events or summary measures
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

Apêndice A - Instrumento de Coleta de Dados

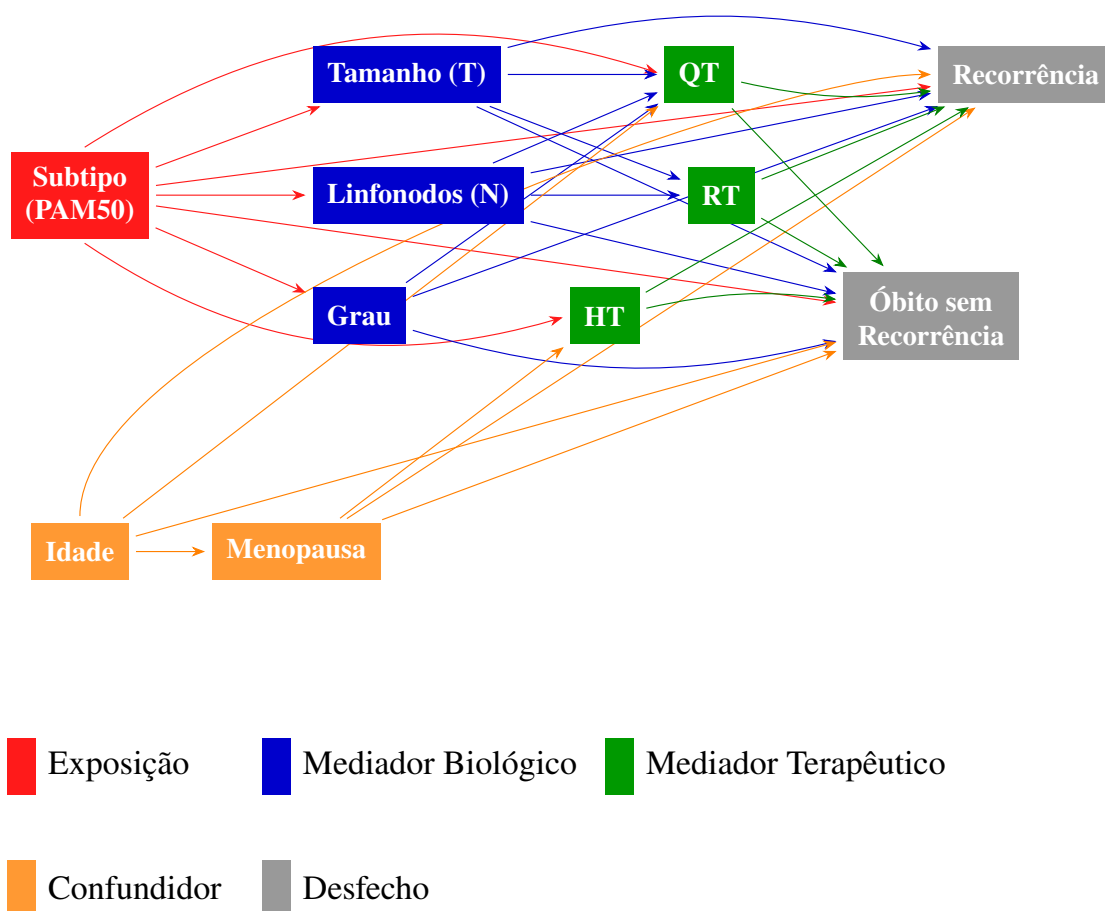
CHECKLIST

BASE DE DADOS:
URL:
Forma de coleta: Download do arquivo
VARIÁVEIS DISPONÍVEIS CONFORME OBJETIVOS DA PESQUISA
Variáveis sociodemográficas: ☐ Idade ☐ Sexo ☐ Escolaridade ☐ Localização geográfica ☐ Raça/cor Variáveis clínicas, histopatológicas e moleculares: ☐ Estadiamento do tumor ☐ Tipo histológico ☐ Grau histológico ☐ Perfil de Receptores Hormonais Variáveis terapêuticas Tipo de tratamento: ☐ Cirúrgico ☐ Quimioterápico ☐ Radioterápico ☐ Hormonal ☐ Outros ☐ Data do diagnóstico ☐ Data do tratamento ☐ Informações sobre recidiva Variáveis de Sobrevida: ☐ Status vital (vivo/morto) ☐ Data do óbito ☐ Data da última informação

Apêndice B - Diagnóstico de suposições e validação do modelo de recorrência do tumor

Este apêndice apresenta as análises suplementares que fundamentaram o modelo de riscos competitivos utilizado na análise de recorrência do tumor.

Figura B1 – Grafo Acíclico Direcionado (DAG): modelo estrutural para avaliação da recorrência e riscos competitivos



Nota: Tamanho (T), Linfonodos (N) e Grau histológico são mediadores biológicos da relação entre subtipo e desfecho, descritos na literatura como preditores independentes de risco. Os tratamentos — QT (Quimioterapia), RT (Radioterapia) e HT (Hormonioterapia) — embora integrem o sistema causal, não são alvo desta análise específica, razão pela qual não foram incluídos como covariáveis de ajuste nos modelos multivariados, a fim de evitar o viés de ajuste excessivo. Devido à colinearidade detectada entre idade e menopausa ($r=0,745$), optou-se pela manutenção exclusiva da idade no modelo final.

Fonte: Dados da pesquisa (2026).

Tabela B1 – Diagnóstico de multicolinearidade dos preditores de recorrência

Preditor	GL	GVIF	GVIF ajustado ^(a)
Tamanho do Tumor	1	1,027	1,013
Número de Linfonodos	1	1,027	1,013
Idade	1	1,085	1,041
Grau Histológico	2	1,317	1,071
Subtipo Molecular (PAM50)	5	1,385	1,033

Notas: GL: Graus de Liberdade; GVIF: *Generalized Variance Inflation Factor*.

^(a) O GVIF Ajustado é calculado pela fórmula $GVIF^{1/(2 \times GL)}$. Esta correção normaliza a métrica em função dos graus de liberdade, permitindo a comparação direta do nível de colinearidade entre variáveis contínuas e categóricas.

Fonte: Dados da pesquisa (2026).

Tabela B2 – Modelo multivariado de riscos proporcionais de Cox causa-específico para o período completo e janelas temporais

Variável	0–120 meses		0–60 meses		61–120 meses	
	HR (IC95%)	p-value	HR (IC95%)	p-value	HR (IC95%)	p-value
Idade (anos)^(a)	0,997 (0,991–1,003)	0,261	0,993 (0,986–1,001)	0,075	1,004 (0,994–1,014)	0,466
Tamanho do tumor (mm)^(a)	1,009 (1,006–1,013)	<0,001	1,010 (1,006–1,014)	<0,001	1,009 (1,001–1,018)	0,035
Linfonodos positivos^(a)	1,061 (1,049–1,073)	<0,001	1,060 (1,045–1,074)	<0,001	1,062 (1,041–1,083)	<0,001
Grau Histológico^(b)						
Grau 2	1,456 (1,038–2,042)	0,030	1,368 (0,843–2,220)	0,204	1,584 (0,988–2,540)	0,056
Grau 3	1,610 (1,140–2,273)	0,007	1,782 (1,096–2,897)	0,019	1,399 (0,852–2,299)	0,185
Subtipo Molecular^(b)						
Luminal B	1,476 (1,205–1,809)	<0,001	1,813 (1,371–2,398)	<0,001	1,207 (0,892–1,635)	0,223
HER2-enriquecido	1,678 (1,309–2,150)	<0,001	2,367 (1,722–3,252)	<0,001	0,971 (0,622–1,517)	0,898
Basal-like	1,317 (0,994–1,744)	0,055	2,407 (1,717–3,374)	<0,001	0,266 (0,126–0,564)	<0,001
Claudin-low	0,957 (0,711–1,286)	0,769	1,242 (0,840–1,836)	0,277	0,728 (0,457–1,159)	0,181
Normal-like	1,415 (1,050–1,907)	0,023	1,547 (1,021–2,344)	0,039	1,371 (0,890–2,112)	0,153

Notas: HR: *Hazard Ratio* (Causa-específico); IC95%: Intervalo de Confiança de 95%.

^(a) Para variáveis contínuas, o HR representa o risco associado ao incremento de uma unidade (1 ano, 1 mm ou 1 linfonodo).

^(b) As categorias de referência para Grau Histológico e Subtipo Molecular foram Grau 1 e Luminal A, respectivamente.

Fonte: Dados da pesquisa (2026).

Tabela B3 – Diagnóstico do pressuposto da proporcionalidade de riscos no modelo de recorrência

Variável	Seguimento completo (0–120 meses)		Precoce (0–60 meses)		Tardia (61–120 meses)	
	Estatística (χ^2)	p	Estatística (χ^2)	p	Estatística (χ^2)	p
Idade	5,68	0,017	0,09	0,768	3,24	0,072
Tamanho	2,90	0,089	0,98	0,323	3,95	0,047
Linfonodos	0,07	0,789	0,10	0,746	0,03	0,872
Grau Histológico	30,82	<0,001	6,57	0,037	0,42	0,810
Subtipo Molecular	81,06	<0,001	28,53	<0,001	8,67	0,123
GLOBAL	85,18	<0,001	34,64	<0,001	14,85	0,138

Nota: Valores em negrito indicam violação da premissa de riscos proporcionais ($p < 0,05$). O teste Global da janela tardia ($p = 0,138$) confirma que, após os 60 meses, o modelo atende aos pressupostos de proporcionalidade.

Fonte: Dados da pesquisa (2026).

Tabela B4 – Distribuição dos eventos por variáveis categóricas e períodos de seguimento

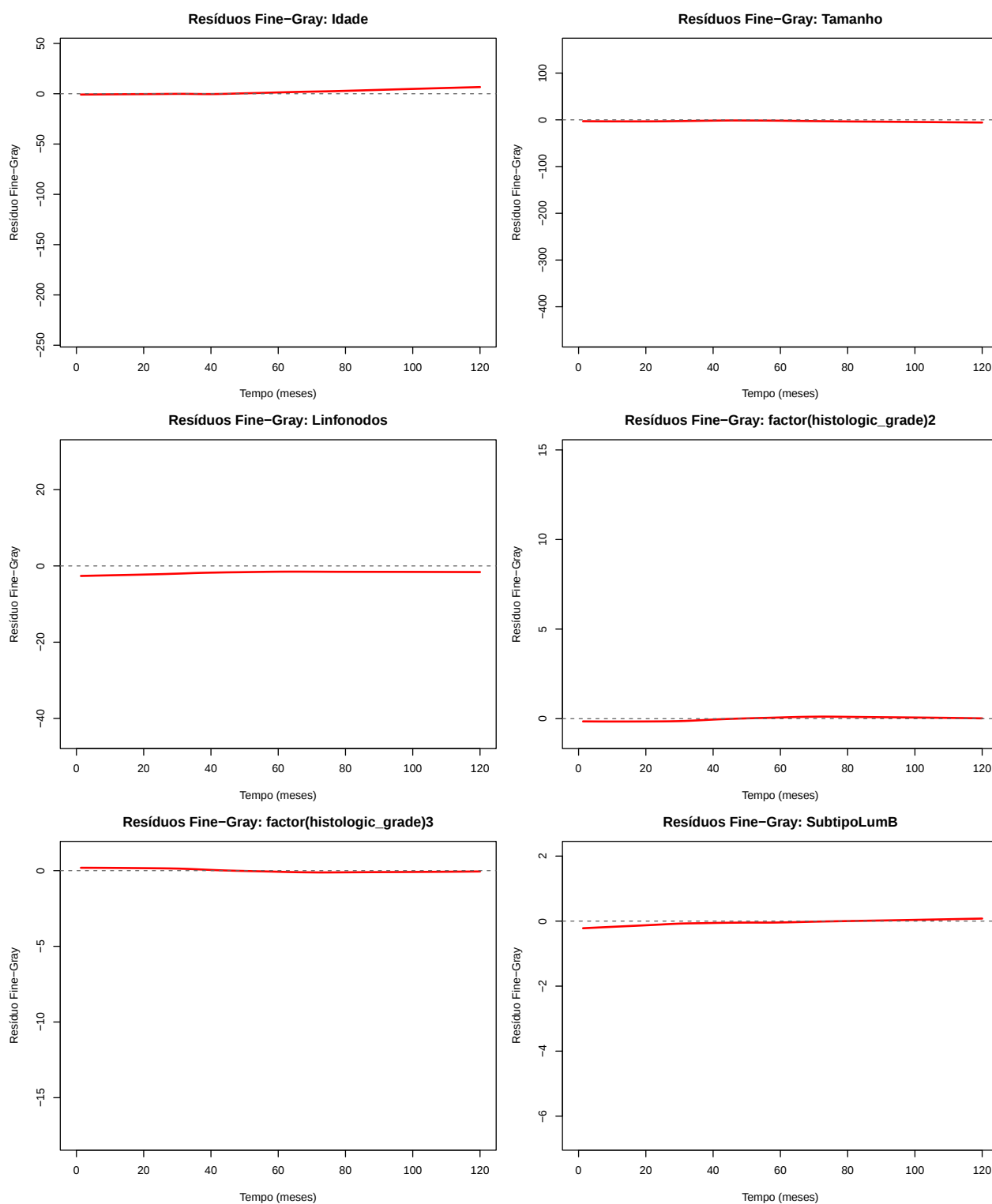
Variável	Precoce (0–60 meses)			Tardio (61–120 meses)		
	Censura	Recorrência	Óbito sem recorrência	Censura	Recorrência	Óbito sem recorrência
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Subtipo Molecular						
<i>Basal-like</i>	112 (58,03)	76 (39,38)	5 (2,59)	88 (80,00)	4 (3,64)	18 (16,36)
HER2-enriquecido	118 (57,28)	81 (39,32)	7 (3,40)	78 (69,03)	20 (17,70)	15 (13,27)
Luminal A	515 (80,47)	93 (14,53)	32 (5,00)	355 (72,01)	57 (11,56)	81 (16,43)
Luminal B	289 (66,13)	129 (29,52)	19 (4,35)	170 (61,82)	57 (20,73)	48 (17,45)
<i>Normal-like</i>	88 (70,40)	30 (24,00)	7 (5,60)	54 (62,79)	16 (18,60)	16 (18,60)
<i>Claudin-low</i>	144 (77,84)	38 (20,54)	3 (1,62)	112 (82,96)	12 (8,89)	11 (8,15)
Grau Histológico						
Grau 1	136 (84,47)	19 (11,80)	6 (3,73)	105 (82,03)	11 (8,59)	12 (9,38)
Grau 2	551 (76,96)	135 (18,85)	30 (4,19)	398 (75,09)	79 (14,91)	53 (10,00)
Grau 3	579 (63,70)	293 (32,23)	37 (4,07)	427 (77,08)	76 (13,72)	51 (9,21)

Nota: Os percentuais na janela 0–60 meses foram calculados sobre o total da amostra original ($n = 1.786$). Os percentuais na janela 61–120 meses foram calculados sobre o total de indivíduos em risco, que atingiram o marco de 5 anos livres de recorrência. ($n = 1.212$). A soma de recorrências, óbitos e censuras em cada linha corresponde ao total de indivíduos que iniciaram a respectiva janela.

Fonte: Dados da pesquisa (2026).

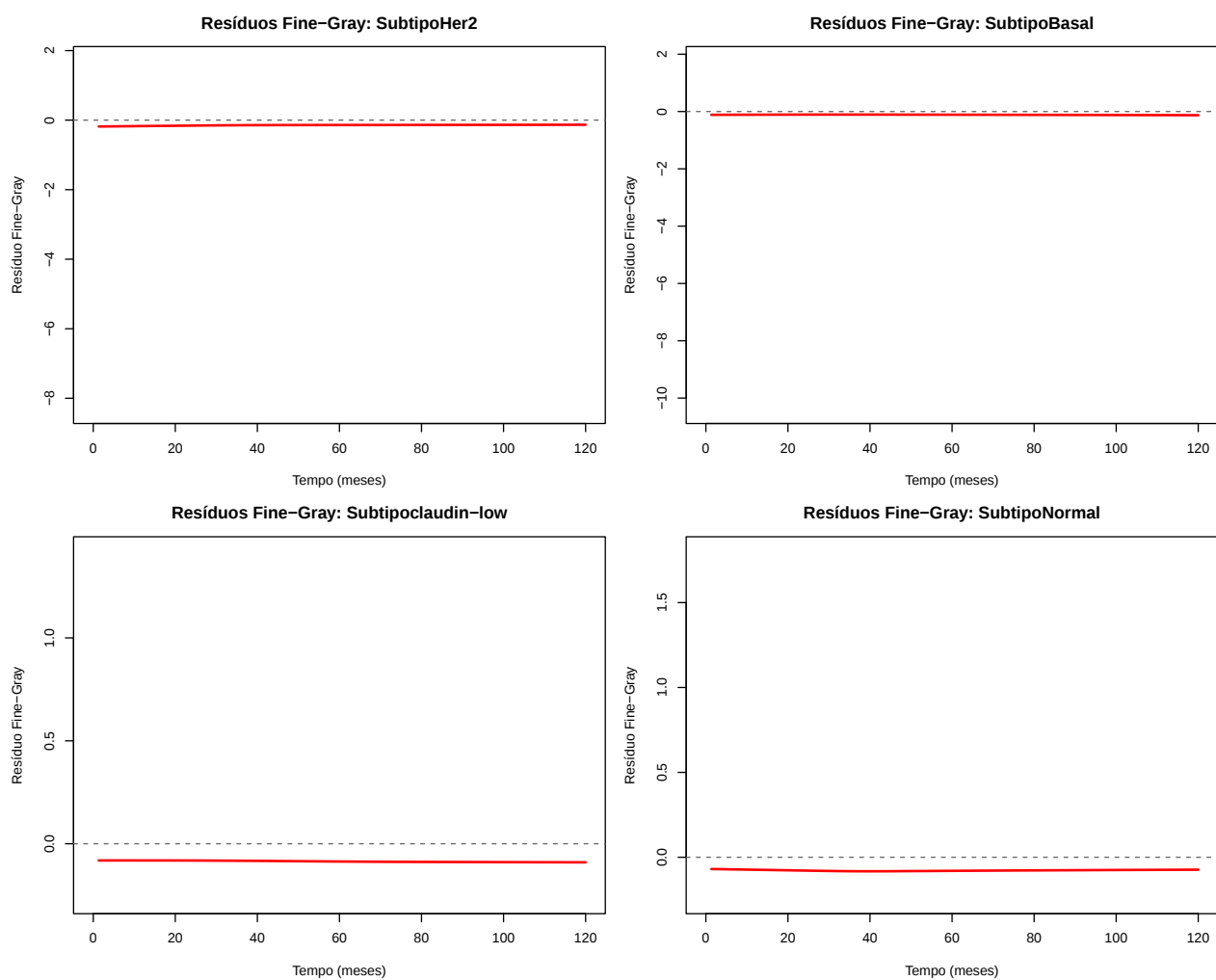
Figura B2 – Resíduos do modelo de recorrência Fine–Gray (0–120 meses)

11



¹¹Os gráficos apresentam os resíduos do modelo de Fine–Gray. A linha vermelha suavizada indica a tendência ao longo do tempo. A linha tracejada em zero (0) representa a referência.

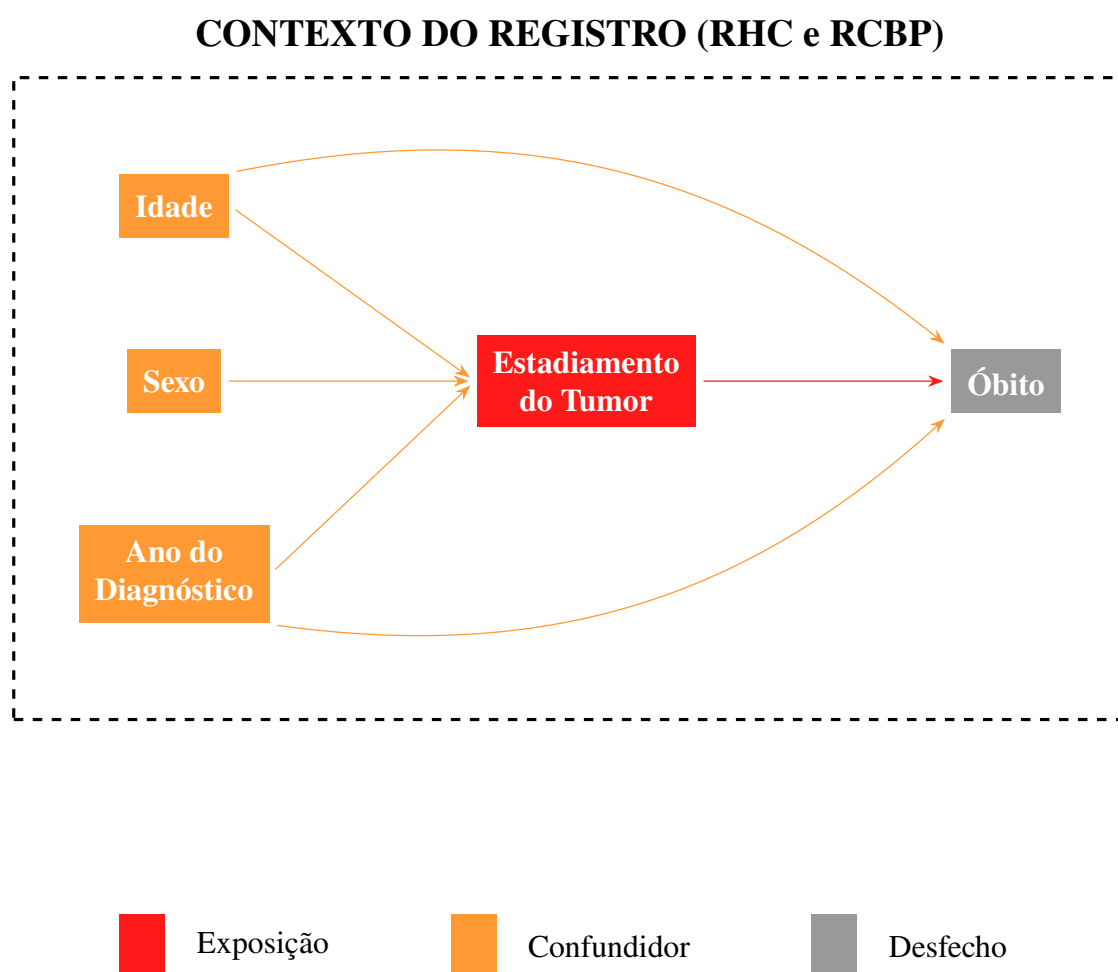
Figura B2 – Resíduos de Schoenfeld do modelo de recorrência (continuação)



Apêndice C - Diagnóstico de suposições e validação do modelo de sobrevivência

Este apêndice apresenta as análises suplementares que fundamentaram a escolha das bases de dados utilizadas nos modelos de risco proporcionais.

Figura C1 – Grafo Acíclico Direcionado (DAG): Modelo estrutural dos fatores clínicos associados ao óbito



Nota: O DAG apresenta o modelo estrutural para a análise de sobrevida em 5 anos. A linha pontilhada indica que o tipo de registro atua como o contexto sob o qual as relações são observadas. As variáveis de idade, sexo e ano do diagnóstico são tratadas como confundidores, pois influenciam simultaneamente a exposição (estadiamento) e o desfecho (óbito), sendo necessário ajustá-las para estimar a associação condicionada entre o estadiamento e a sobrevida em cada sistema.

Fonte: Dados da pesquisa (2026).

Tabela C1 – Diagnóstico de multicolinearidade dos preditores de sobrevida

Preditor	GL	GVIF	GVIF ajustado ^(a)
RHC Estadual			
Sexo	1	1,000	1,000
Estadiamento	3	1,006	1,001
Ano diagnóstico	1	1,006	1,003
RCBP Internacional			
Sexo	1	1,001	1,000
Estadiamento	3	1,002	1,000
Ano diagnóstico	1	1,001	1,001

Notas: GL: Graus de Liberdade; GVIF: *Generalized Variance Inflation Factor*.

^(a) O GVIF Ajustado é calculado pela fórmula $GVIF^1/(2 \times GL)$. Esta correção normaliza a métrica em função dos graus de liberdade, permitindo a comparação direta do nível de colinearidade entre variáveis contínuas e categóricas.

Fonte: Dados da pesquisa (2026).

Tabela C2 – Análise de Sensibilidade: Comparação de Hazard Ratios (IC 95%) para o RHC Estadual.

Variável	Exclusão HR (IC 95%)	“Ausentes” HR (IC 95%)
Estadiamento		
I (Referência)		
II	2,199 (2,081–2,324)	2,200 (2,082–2,325)
III	6,399 (6,071–6,745)	6,397 (6,069–6,744)
IV	19,423 (18,384–20,521)	19,367 (18,331–20,462)
“Ausentes” ^(a)	—	4,693 (4,240–5,194)
Ajustes		
Idade (anos)	1,015 (1,014–1,016)	1,015 (1,014–1,016)
Sexo (Masculino)	1,061 (0,914–1,233)	1,064 (0,919–1,233)
Ano Diagnóstico	0,976 (0,974–0,979)	0,977 (0,974–0,980)

Nota: HR: Razão de Riscos (*Hazard Ratios*); IC 95%: intervalo de confiança de 95%. Modelos multivariados de riscos proporcionais de Cox para sobrevida global em 5 anos. Número total de observações: Modelo de Exclusão ($N = 90.192$; Eventos = 23.752), “Ausentes” ($N = 91.850$; Eventos = 24.239).

^(a) Embora a categoria “Ausentes” apresente risco, a sua inclusão no modelo não alterou de forma relevante os coeficientes das demais variáveis, sugerindo a robustez do modelo de exclusão.

Fonte: Dados da pesquisa (2026).

Tabela C3 – Análise de Sensibilidade: Comparação de Hazard Ratios (IC 95%) para o RCBP Internacional.

Variável	Exclusão HR (IC 95%)	“Ausentes” HR (IC 95%)
Estadiamento		
I (Referência)		
II	2,005 (1,975–2,035)	2,003 (1,973–2,034)
III	5,132 (5,049–5,217)	5,125 (5,042–5,210)
IV	15,021 (14,661–15,390)	14,972 (14,613–15,339)
“Ausentes” ^(a)	—	2,731 (2,657–2,808)
Ajustes		
Idade (anos)	1,050 (1,049–1,051)	1,050 (1,049–1,051)
Sexo (Masculino)	1,287 (1,222–1,356)	1,300 (1,236–1,368)
Ano Diagnóstico	0,973 (0,972–0,975)	0,975 (0,974–0,976)

Nota: HR: Razão de Riscos (*Hazard Ratios*); IC 95%: intervalo de confiança de 95%. Modelos multivariados de riscos proporcionais de Cox para sobrevida global truncada a 60 meses. Número total de observações: Modelo de Exclusão ($N = 719.620$; Eventos = 103.991), “Ausentes” ($N = 746.567$; Eventos = 110.037).

^(a) Embora a categoria “Ausentes” apresente risco, a sua inclusão no modelo não alterou de forma relevante os coeficientes das demais variáveis, sugerindo a robustez do modelo de exclusão.

Fonte: Dados da pesquisa (2026).

Tabela C4 – Diagnóstico do pressuposto da proporcionalidade de riscos nos modelos de sobrevida global

Variável	RHC Estadual		RCBP Internacional	
	Estatística (χ^2)	p	Estatística (χ^2)	p
Idade	7,33	0,007	5,60	0,018
Sexo	0,07	0,791	0,75	0,388
Estadiamento	414,60	<0,001	926,17	<0,001
Ano diagnóstico	0,10	0,749	68,26	<0,001
GLOBAL	422,14	<0,001	983,77	<0,001

Nota: Valores em negrito indicam violação da premissa de riscos proporcionais ($p < 0,05$).

Fonte: Dados da pesquisa (2026).

Tabela C5 – Diagnóstico do pressuposto da proporcionalidade de riscos nos modelos de sobrevida específica

Variável	RHC Estadual		RCBP Internacional	
	Estatística (χ^2)	p	Estatística (χ^2)	p
Idade	98,90	<0,001	415,58	<0,001
Sexo	0,93	0,340	7,03	0,008
Estadiamento	393,00	<0,001	787,07	<0,001
Ano diagnóstico	0,00	0,990	30,16	<0,001
GLOBAL	489,00	<0,001	1341,65	<0,001

Nota: Valores em negrito indicam violação da premissa de riscos proporcionais ($p < 0,05$).

Fonte: Dados da pesquisa (2026).

Tabela C6 – Análise de sobrevida: contagem de eventos (óbitos) por tipo de registro

Indicador	RHC Estadual	RCBP Internacional
Total óbitos (todas causas)	39.572	257.852
Total óbitos (por câncer de mama)	25.052	100.545
Óbitos até 60 meses (todas causas)	23.752	103.991
Óbitos até 60 meses (por câncer de mama)	17.559	55.201
Óbitos até 60 meses (por outras causas)	6.193	48.790

Nota: A contagem exclui os dados com informações ausentes para as variáveis do modelo.

Fonte: Dados da pesquisa (2026).

Figura C2 – RHC Estadual: Resíduos do modelo de sobrevida global (Cox)

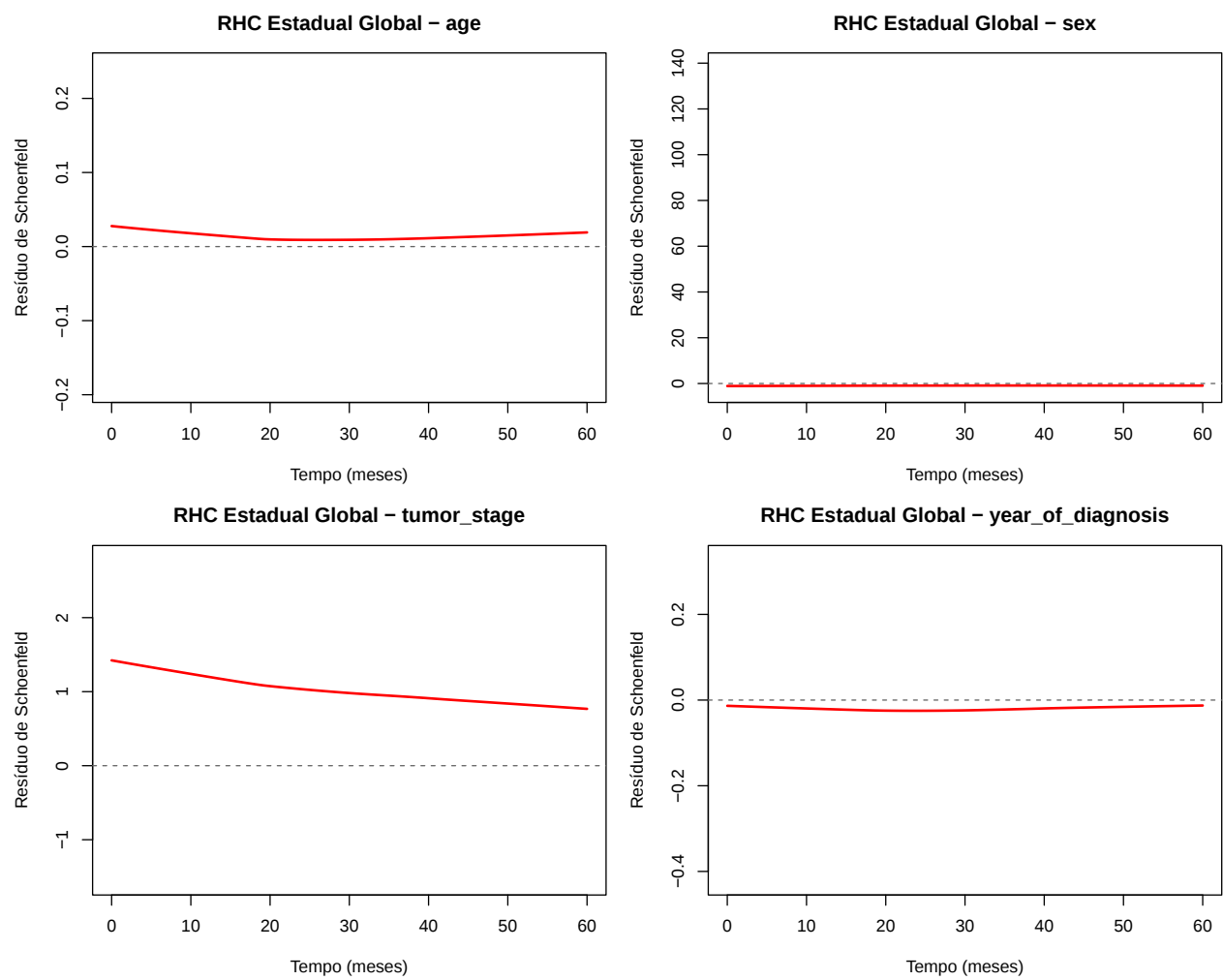


Figura C3 – RHC Estadual: Resíduos do modelo de sobrevida específica (Cox)

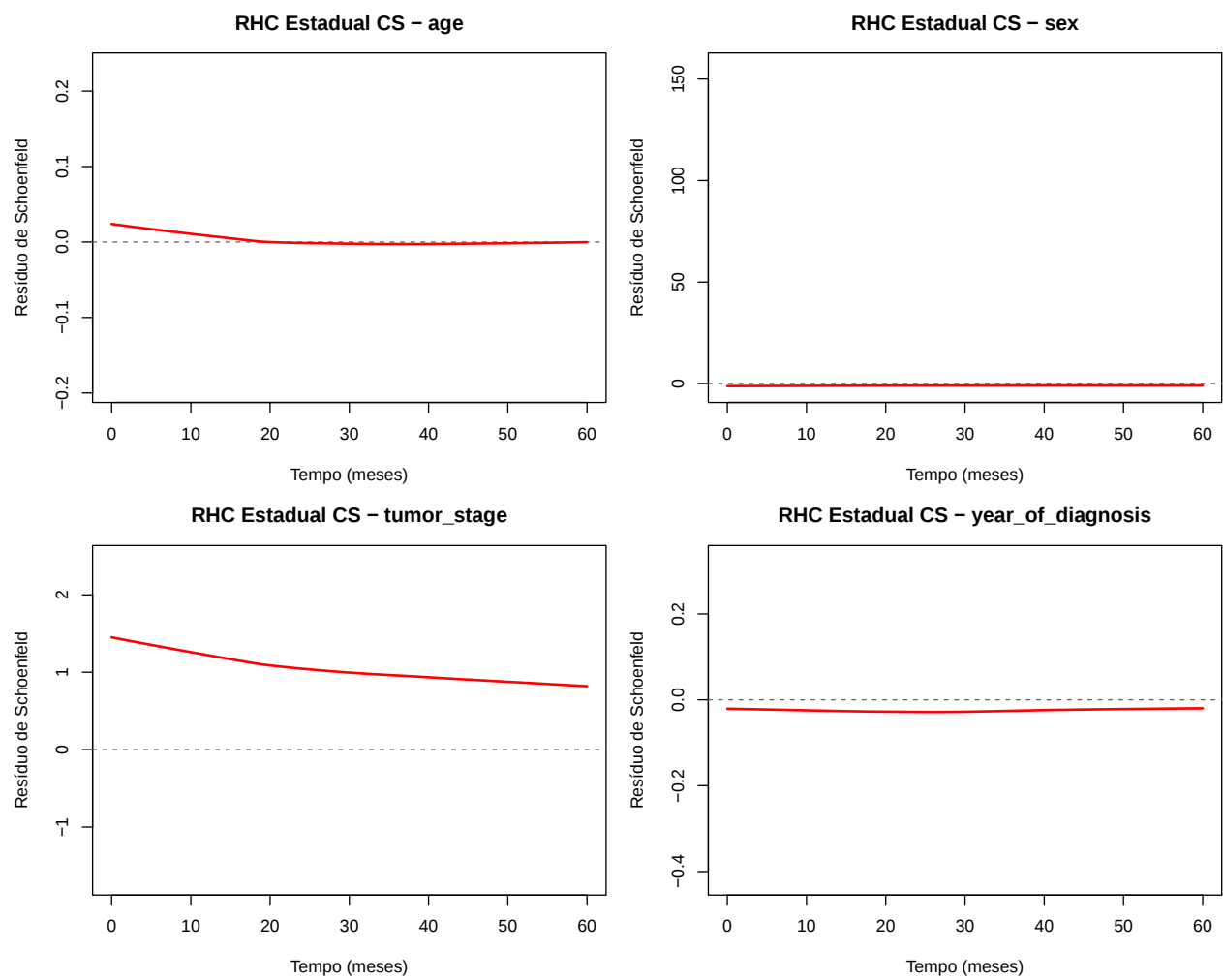


Figura C4 – RCBP Internacional: Resíduos do modelo de sobrevida global (Cox)

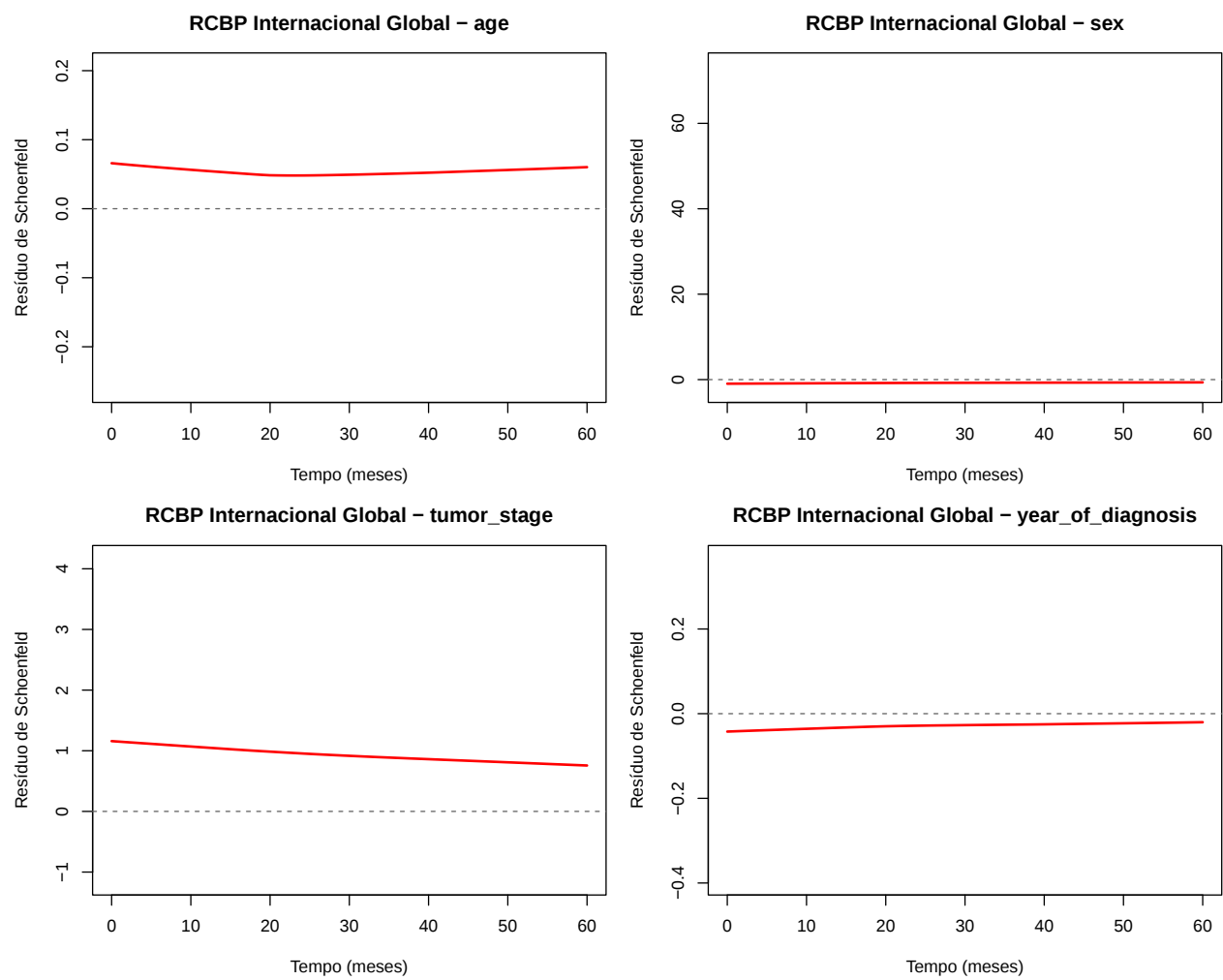
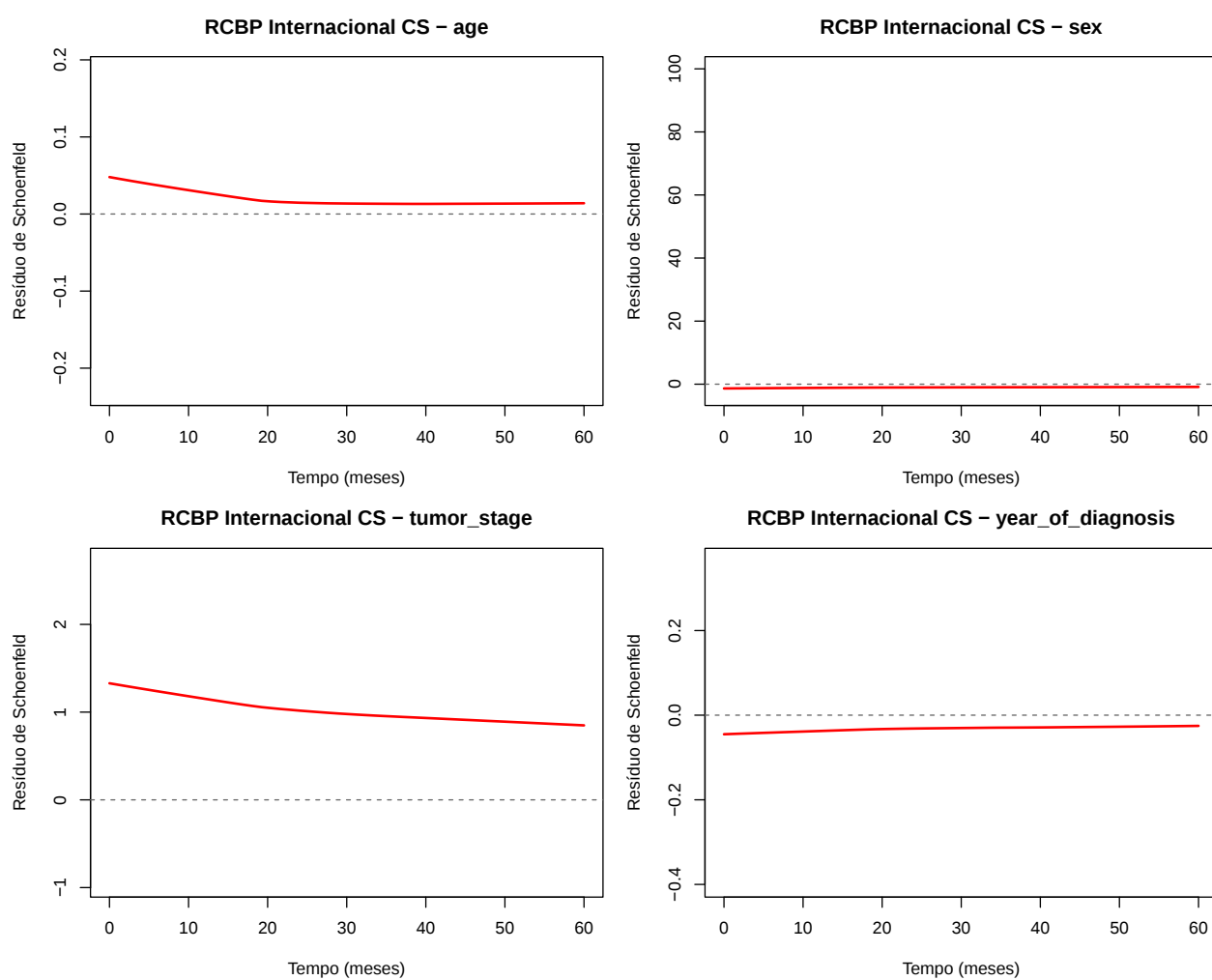


Figura C5 – RCBP Internacional: Resíduos do modelo de sobrevida específica (Cox)



Apêndice D - Tratamento do câncer de mama no RHC Nacional

Tabela D1 – Frequência das modalidades terapêuticas segundo a região de tratamento no RHC Nacional

Região e Modalidade	Sim n (%)	Não n (%)	Ignorado n (%)
Cirurgia			
Centro-Oeste	5.443 (56,25%)	4.209 (43,49%)	25 (0,26%)
Nordeste	43.798 (62,32%)	26.378 (37,53%)	99 (0,14%)
Norte	6.490 (57,67%)	4.689 (41,67%)	75 (0,67%)
Sudeste	114.724 (75,66%)	36.712 (24,21%)	188 (0,12%)
Sul	28.534 (51,81%)	26.376 (47,90%)	161 (0,29%)
Quimioterapia			
Centro-Oeste	5.779 (63,67%)	3.873 (42,67%)	25 (0,28%)
Nordeste	48.548 (69,08%)	21.628 (30,78%)	99 (0,14%)
Norte	7.826 (69,54%)	3.353 (29,80%)	75 (0,67%)
Sudeste	104.121 (68,67%)	47.315 (31,21%)	188 (0,12%)
Sul	31.425 (57,06%)	23.485 (42,65%)	161 (0,29%)
Radioterapia			
Centro-Oeste	4.518 (49,77%)	5.134 (56,56%)	25 (0,28%)
Nordeste	35.398 (50,37%)	34.778 (49,49%)	99 (0,14%)
Norte	6.089 (54,11%)	5.090 (45,23%)	75 (0,67%)
Sudeste	74.486 (49,13%)	76.950 (50,75%)	188 (0,12%)
Sul	23.175 (42,08%)	31.735 (57,63%)	161 (0,29%)
Hormonioterapia			
Centro-Oeste	2.590 (28,53%)	7.062 (77,80%)	25 (0,28%)
Nordeste	25.859 (36,80%)	44.317 (63,06%)	99 (0,14%)
Norte	3.066 (27,24%)	8.113 (72,09%)	75 (0,67%)
Sudeste	79.602 (52,50%)	71.834 (47,38%)	188 (0,12%)
Sul	17.331 (31,47%)	37.580 (68,24%)	160 (0,29%)
Imunoterapia			
Centro-Oeste	3 (0,03%)	9.649 (106,30%)	25 (0,28%)
Nordeste	555 (0,79%)	69.621 (99,07%)	99 (0,14%)
Norte	24 (0,21%)	11.155 (99,12%)	75 (0,67%)
Sudeste	877 (0,58%)	150.559 (99,30%)	188 (0,12%)
Sul	134 (0,24%)	54.776 (99,46%)	161 (0,29%)

Nota: Total de registros da amostra: $N = 297.901$. Os percentuais foram calculados em relação ao total de registros de cada região de tratamento. Como um mesmo indivíduo pode ter recebido múltiplas modalidades terapêuticas (terapias combinadas), as frequências apresentadas para as diferentes modalidades não são mutuamente exclusivas e não devem ser somadas entre si. Os percentuais devem ser interpretados como a proporção de indivíduos da respectiva região de tratamento que apresentaram registro da modalidade terapêutica correspondente.

Fonte: Dados da pesquisa (2026).

Tabela D2 – Distribuição dos deslocamentos para tratamento oncológico fora da região de residência, por modalidade e segundo as regiões e faixas de distância (2000–2015)

Região e Modalidade	Até 80 km <i>n</i> (%)	80–250 km <i>n</i> (%)	Mais de 250 km <i>n</i> (%)
Região Norte			
Cirurgia	376 (13,31%)	988 (34,96%)	1.462 (51,73%)
Quimioterapia	508 (14,81%)	1.228 (35,80%)	1.694 (49,39%)
Radioterapia	412 (16,14%)	803 (31,46%)	1.338 (52,41%)
Hormonioterapia	182 (16,61%)	364 (33,21%)	550 (50,18%)
Imunoterapia	2 (14,29%)	6 (42,86%)	6 (42,86%)
Região Nordeste			
Cirurgia	8.579 (33,44%)	8.734 (34,04%)	8.345 (32,52%)
Quimioterapia	9.955 (34,79%)	10.213 (35,69%)	8.449 (29,52%)
Radioterapia	6.806 (33,95%)	7.062 (35,23%)	6.179 (30,82%)
Hormonioterapia	5.329 (36,70%)	4.839 (33,32%)	4.354 (29,98%)
Imunoterapia	83 (36,24%)	76 (33,19%)	70 (30,57%)
Região Centro-Oeste			
Cirurgia	360 (17,89%)	758 (37,67%)	894 (44,43%)
Quimioterapia	442 (19,31%)	857 (37,44%)	990 (43,25%)
Radioterapia	302 (17,47%)	621 (35,92%)	806 (46,62%)
Hormonioterapia	132 (17,12%)	371 (48,12%)	268 (34,76%)
Imunoterapia	0 (0,00%)	0 (0,00%)	1 (100,00%)
Região Sudeste			
Cirurgia	32.367 (58,36%)	19.156 (34,54%)	3.942 (7,11%)
Quimioterapia	30.685 (58,30%)	18.270 (34,71%)	3.676 (6,98%)
Radioterapia	21.035 (56,30%)	13.737 (36,77%)	2.588 (6,93%)
Hormonioterapia	22.553 (58,71%)	13.402 (34,89%)	2.457 (6,40%)
Imunoterapia	245 (45,97%)	218 (40,90%)	70 (13,13%)
Região Sul			
Cirurgia	8.647 (51,66%)	6.880 (41,10%)	1.212 (7,24%)
Quimioterapia	9.526 (52,08%)	7.739 (42,31%)	1.026 (5,61%)
Radioterapia	6.912 (49,23%)	6.156 (43,84%)	973 (6,93%)
Hormonioterapia	5.500 (56,16%)	3.715 (37,93%)	579 (5,91%)
Imunoterapia	38 (50,00%)	24 (31,58%)	14 (18,42%)

Nota: Valores apresentados como *n* (%). A tabela foi construída com base em um total de 155.186 registros de indivíduos que realizaram tratamento fora do município de residência. Como cada indivíduo pode ter realizado mais de uma modalidade terapêutica (terapias combinadas), as somas de casos por modalidade e região podem ultrapassar o *n* amostral. Os percentuais foram calculados sobre o total de casos de cada modalidade dentro de sua respectiva região. Como algumas modalidades apresentam número reduzido de registros em determinadas regiões (por exemplo, imunoterapia no Norte e Centro-Oeste), os percentuais nessas categorias devem ser interpretados com cautela, pois refletem pequenas contagens absolutas. Os resultados devem ser entendidos como *frequências relativas de deslocamento* dentro de cada modalidade e região, e não como distribuições exclusivas por paciente.

Fonte: Dados da pesquisa (2026).

Apêndice E - Declaração de Uso de Inteligência Artificial Generativa (IAG)

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA (IAG)

Em atendimento às diretrizes da **Resolução COLPPGIPA Nº 11, de 02 de junho de 2026**, que dispõe sobre o uso ético, transparente e responsável de ferramentas de Inteligência Artificial Generativa (IAG) no âmbito do Programa de Pós-Graduação em Imunologia e Parasitologia Aplicadas (PPGIPA) da Universidade Federal de Uberlândia (UFU), eu, **Gisele Cristiane da Silva Dias**, autora da dissertação intitulada **“Impacto dos subtipos moleculares e variáveis clínico-patológicas nos desfechos do câncer de mama: um estudo entre grandes bases de dados”**, declaro que, durante o desenvolvimento deste trabalho, as seguintes ferramentas de IAG foram utilizadas:

- Copilot
- Gemini
- Grammarly

As referidas ferramentas foram empregadas estritamente para os seguintes propósitos:

- Aprimoramento da gramática, ortografia e clareza textual da redação;
- Tradução e revisão do resumo para a língua inglesa (*Abstract*);
- Formatação e estruturação automatizada do código LaTeX de tabelas a partir de dados previamente calculados e analisados.

Declaro que o uso das ferramentas de IAG seguiu os preceitos éticos e de integridade científica, não englobando a criação, manipulação ou fabricação de dados, imagens ou figuras. Toda a concepção, validação dos resultados e redação do trabalho são de minha autoria e responsabilidade.

Documento assinado digitalmente
GISELE CRISTIANE DA SILVA DIAS
Data: 08/08/2026 09:39:26-0300
Verifique em <https://validar.itl.gov.br>

Gisele Cristiane da Silva Dias
Discente do PPGIPA/UFU